

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2000-47699

(P2000-47699A)

(43)公開日 平成12年2月18日(2000.2.18)

| | | | |
|--------------------------|------|-------------------------------|-------------|
| (51)Int.Cl. ⁷ | 識別記号 | F I | マークシート*(参考) |
| G 1 0 L 21/02 | | G 1 0 L 9/00 | F 5 D 0 1 5 |
| 19/00 | | | H 5 J 0 2 3 |
| G 0 1 S 3/808 | | G 0 1 S 3/808 | 5 J 0 8 3 |
| 3/86 | | 3/86 | |
| H 0 3 H 21/00 | | G 1 0 L 3/02 | 3 0 1 E |
| | | 3 0 1 E | |
| | | 審査請求 未請求 請求項の数 8 O L (全 28 頁) | 最終頁に続く |

(21)出願番号 特願平10-217519

(22)出願日 平成10年7月31日(1998.7.31)

(71)出願人 000003078

株式会社東芝

神奈川県川崎市幸区堀川町72番地

(72)発明者 金澤 博史

兵庫県神戸市東灘区本山南町8丁目6番26号 株式会社東芝関西研究所内

(72)発明者 赤嶺 政巳

兵庫県神戸市東灘区本山南町8丁目6番26号 株式会社東芝関西研究所内

(74)代理人 100058479

弁理士 鈴江 武彦 (外6名)

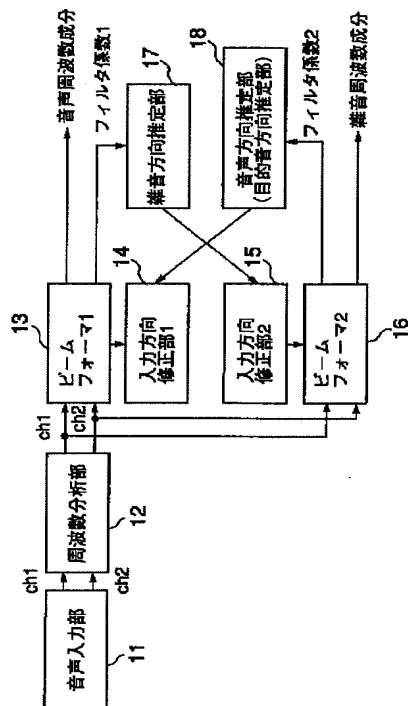
最終頁に続く

(54)【発明の名称】 雑音抑圧処理装置および雑音抑圧処理方法

(57)【要約】

【課題】 計算量を大幅に削減できる雑音抑圧処理を可能にすること。

【解決手段】 話者音声異なる複数位置で検出する手段11と、この受信位置別音声信号毎に周波数分析して各チャンネル(ch)別の周波数成分を得る分析手段12と、この各chの周波数成分について所望方向外の感度が低くなるように計算したフィルタ(F)係数によるフィルタ(F)処理により話者方向の雑音を抑圧し、目的音声成分を得る第1ビームフォーマ処理手段(B1)13と、分析手段にて得た複数chの周波数成分について所望方向外を低感度化するF処理により話者音声を抑圧し、雑音成分を得る第2ビームフォーマ処理手段(B2)16と、B1のF係数から雑音方向を推定し、B2のF係数から目的音方向を推定する推定手段17,18と、B1において入力対象となる目的音の到来方向である第1入力方向を、推定手段の推定目的音方向に基づき修正し、B2において入力対象とする雑音の到来方向である第2入力方向を、推定手段の推定雑音方向に基づき修正する修正手段14,15を具備する。



【特許請求の範囲】

【請求項 1】話者の発声した音声を 2 箇所以上の異なった位置で受音する音声入力手段と、
前記受音位置に対応する音声信号のチャンネルごとに周波数分析を行って複数チャンネルの周波数成分を出力する周波数分析手段と、
前記複数チャンネルの周波数成分を用いて適応フィルタ処理により目的の音声以外の到来雑音の抑圧処理を行って目的の音声を出力する第 1 のビームフォーマ処理手段と、
前記複数チャンネルの周波数成分を用いて適応フィルタ処理により目的の音声の抑圧処理を行って雑音を出力する第 2 のビームフォーマ処理手段と、
前記第 1 のビームフォーマ処理手段で計算されるフィルタ係数から雑音方向を推定する雑音方向推定手段と、
前記第 2 のビームフォーマ処理手段で計算されるフィルタ係数から目的音方向を推定する目的音方向推定手段と、
前記第 1 のビームフォーマにおいて入力対象となる目的音の到来方向である第 1 の入力方向を、前記目的音方向推定手段で推定された目的音方向に基づいて逐次修正する目的音方向修正手段と、
前記第 2 のビームフォーマにおいて入力対象とする雑音の到来方向である第 2 の入力方向を、前記雑音方向推定手段で推定された雑音方向に基づいて逐次修正する雑音方向修正手段とを具備し、
逐次、音声周波数成分と雑音周波数成分とを別々に出力することを特徴とする雑音抑圧装置。

【請求項 2】話者の発声した音声を少なくとも異なる 2 箇所以上の位置で受音する音声入力手段と、前記受音位置に対応する音声信号のチャンネル毎に周波数分析を行って複数チャンネルの周波数成分を出力する周波数分析手段と、この周波数分析手段にて得られる前記複数チャンネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声以外の音声を抑圧する到来雑音抑圧処理を行い、目的音声成分を得る第 1 のビームフォーマ処理手段と、
前記周波数分析手段にて得られる前記複数チャンネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声を抑圧し、第 1 の雑音成分を得る第 2 のビームフォーマ処理手段と、
前記周波数分析手段にて得られる前記複数チャンネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声を抑圧し、第 2 の雑音成分を得る第 2 のビームフォーマ処理手段と、
前記第 1 のビームフォーマ処理手段で計算されるフィルタ係数から雑音方向を推定する雑音方向推定手段と、
前記第 2 のビームフォーマ処理手段で計算されるフィル

タ係数から第 1 の目的音方向を推定する第 1 の目的音方向推定手段と、

前記第 3 の適応ビームフォーマ処理手段で計算されるフィルタ係数から第 2 の目的音方向を推定する第 2 の目的音方向推定手段と、

前記第 1 のビームフォーマにおいて入力対象とする目的音の到来方向である第 1 の入力方向を、前記第 1 の目的音方向推定手段で推定された第 1 の目的音方向と、第 2 の目的音方向推定手段で推定された第 2 の目的音方向のいずれか一方または両方に基づいて逐次修正する第 1 の入力方向修正手段と、

前記雑音方向修正手段で推定された雑音方向が所定の第 1 の範囲にある場合に、前記第 2 のビームフォーマにおいて入力対象とする雑音の到来方向である第 2 の入力方向を該雑音方向に基づいて逐次修正する第 2 の入力方向修正手段と、

前記雑音方向修正手段で推定された雑音方向が所定の第 2 の範囲にある場合に、前記第 3 のビームフォーマにおいて入力対象とする雑音の到来方向である第 3 の入力方向を該雑音方向に基づいて逐次修正する第 3 の入力方向修正手段と、

前記雑音方向推定手段で推定された雑音方向が所定の第 1 の範囲から到来したか所定の第 2 の範囲から到来したかに基づいて前記第 1 および第 2 の出力雑音のいずれか一方を真の雑音出力と決定していずれか一方の雑音を出力すると同時に、第 1 の音声方向推定手段と第 2 の音声方向推定手段のいずれの推定結果が有効であるかを決定していずれか一方の音声方向推定結果を第 1 の入力方向修正手段へ出力する有効雑音決定手段と、

を具備し、逐次、音声周波数成分と雑音周波数成分とを別々に出力することを特徴とする雑音抑圧装置。

【請求項 3】請求項 1 または 2 いずれか 1 項に記載の雑音抑圧装置において、

前記得られた音声周波数を、周波数帯域毎に分割して帯域毎の音声パワーを計算する音声帯域パワー計算手段と、

前記得られた雑音周波数成分を、周波数帯域毎に分割して帯域毎の雑音パワーを計算する雑音帯域パワー計算手段と、

前記音声帯域パワー計算手段と雑音帯域パワー計算手段とから得られる音声と雑音の周波数帯域パワーに基き、音声信号の周波数帯域毎に重みをかけて背景雑音を抑圧するスペクトル減算手段と、からなるスペクトル減算雑音抑圧手段をさらに具備することを特徴とする雑音抑圧装置。

【請求項 4】請求項 1 または 2 いずれか 1 項に記載の雑音抑圧装置において、

前記得られた音声周波数を、周波数帯域毎に分割して帯域毎の音声パワーを計算する音声帯域パワー計算手段と、

前記得られた雑音周波数成分を、周波数帯域毎に分割して帯域毎の雑音パワーを計算する雑音帯域パワー計算手段と、

前記音声入力手段から得られた入力信号を周波数分析した入力信号の周波数成分を周波数帯域毎に分割し、帯域毎の入力パワーを計算する入力帯域パワー計算手段と、前記入力帯域パワーと音声帯域パワーと雑音帯域パワーとに基づき、音声信号の周波数帯域毎に重みをかけて背景雑音を抑圧する修正スペクトル減算手段を具備することを特徴とする雑音抑圧装置。

【請求項 5】話者の発声した音声を 2 箇所以上の異なる位置でそれぞれ受音してそれぞれ別チャンネルの音声信号として得るステップと、

この各チャンネル毎の音声信号を周波数分析してそれぞれチャンネル別に周波数スペクトル成分を得る周波数分析ステップと、

周波数分析ステップにて得られた各チャンネルの周波数成分を用いて適応フィルタ処理により目的の音声以外の到来雑音を抑圧処理し、目的音声を得る第 1 のビームフォーマ処理ステップと、

前記各チャンネルの周波数成分を用いて適応フィルタ処理により目的の音声の抑圧処理を行って雑音成分を得る第 2 のビームフォーマ処理ステップと、

前記第 1 のビームフォーマ処理ステップで計算される適応フィルタで使用したフィルタ係数から雑音方向を推定する雑音方向推定ステップと、

前記第 2 のビームフォーマ処理ステップで計算される適応フィルタで使用したフィルタ係数から目的音方向を推定する目的音方向推定ステップと、

前記第 1 のビームフォーマ処理ステップにおいて入力対象となる目的音の到来方向である第 1 の入力方向を、前記目的音方向推定手段で推定された目的音方向に基づいて逐次修正する目的音方向修正ステップと、

前記第 2 のビームフォーマ処理ステップにおいて入力対象とする雑音の到来方向である第 2 の入力方向を、前記雑音方向推定ステップで推定された雑音方向に基づいて逐次修正する雑音方向修正ステップとを具備し、

逐次、音声周波数成分と雑音周波数成分とを別々に求めることを特徴とする雑音抑圧方法。

【請求項 6】話者の発声した音声を 2 箇所以上の異なる位置でそれぞれ受音してそれぞれ別チャンネルの音声信号として得るステップと、

この各チャンネル毎の音声信号を周波数分析してそれぞれチャンネル別に周波数スペクトル成分を得る周波数分析ステップと、

周波数分析ステップにて得られた各チャンネルの周波数成分を用いて前記複数チャンネルの周波数成分について、所望方向外の感度が低くなるようにしたフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声以外の音声を抑圧する到来雑音抑圧処理を行

い、目的音声成分を得る第 1 のビームフォーマ処理ステップと、

周波数分析ステップにて得られた各チャンネルの周波数成分を用いて前記複数チャンネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声を抑圧し、第 1 の雑音成分を得る第 2 のビームフォーマ処理ステップと、

10 周波数分析ステップにて得られた各チャンネルの周波数成分を用いて前記複数チャンネルの周波数成分について、所望方向外の感度が低くなるようにしたフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声を抑圧し、第 2 の雑音成分を得る第 2 のビームフォーマ処理ステップと、

前記第 1 のビームフォーマ処理ステップで計算されるフィルタ係数から雑音方向を推定する雑音方向推定ステップと、

20 前記第 2 のビームフォーマ処理ステップで計算されるフィルタ係数から第 1 の目的音方向を推定する第 1 の目的音方向推定ステップと、

前記第 3 の適応ビームフォーマ処理ステップで計算されるフィルタ係数から第 2 の目的音方向を推定する第 2 の目的音方向推定ステップと、

前記第 1 のビームフォーマにおいて入力対象とする目的音の到来方向である第 1 の入力方向を、前記第 1 の目的音方向推定手段で推定された第 1 の目的音方向と、第 2 の目的音方向推定手段で推定された第 2 の目的音方向のいずれか一方または両方に基づいて逐次修正する第 1 の入力方向修正ステップと、

30 前記雑音方向修正ステップで推定された雑音方向が所定の第 1 の範囲にある場合に、前記第 2 のビームフォーマ処理ステップにおいて入力対象とする雑音の到来方向である第 2 の入力方向を該雑音方向に基づいて逐次修正する第 2 の入力方向修正ステップと、

前記雑音方向修正ステップで推定された雑音方向が所定の第 2 の範囲にある場合に、前記第 3 のビームフォーマ処理ステップにおいて入力対象とする雑音の到来方向である第 3 の入力方向を該雑音方向に基づいて逐次修正する第 3 の入力方向修正ステップと、

40 前記雑音方向推定ステップで推定された雑音方向が所定の第 1 の範囲から到来したか所定の第 2 の範囲から到来したかに基づいて前記第 1 および第 2 の出力雑音のいずれか一方を真の雑音出力と決定していずれか一方の雑音を出力すると同時に、第 1 の音声方向推定手段と第 2 の音声方向推定手段のいずれの推定結果が有効であるかを決定していずれか一方の音声方向推定結果を第 1 の入力方向修正ステップで使用する音声方向推定結果として与える有効雑音決定ステップと、を具備することを特徴とする雑音抑圧方法。

50 【請求項 7】請求項 5 または 6 いずれか 1 項に記載の雑

音抑圧方法において、

前記得られた音声周波数を、周波数帯域毎に分割して帯域毎の音声パワーを計算する音声帯域パワー計算ステップと、

前記得られた雑音周波数成分を、周波数帯域毎に分割して帯域毎の雑音パワーを計算する雑音帯域パワー計算ステップと、

前記音声帯域パワー計算ステップにて得られる音声の周波数帯域パワーと、雑音帯域パワー計算ステップにて得られる雑音の周波数帯域パワーとに基き、音声信号の周波数帯域毎に重みをかけて背景雑音を抑圧するスペクトル減算ステップと、をさらに具備することを特徴とする雑音抑圧方法。

【請求項 8】請求項 5 または 6 いずれか 1 項に記載の雑音抑圧方法において、

前記得られた音声周波数を、周波数帯域毎に分割して帯域毎の音声パワーを計算する音声帯域パワー計算ステップと、

前記得られた雑音周波数成分を、周波数帯域毎に分割して帯域毎の雑音パワーを計算する雑音帯域パワー計算ステップと、

前記周波数分析ステップにて得られた入力信号の周波数スペクトル成分を周波数帯域毎に分割し、帯域毎の入力パワーを計算する入力帯域パワー計算ステップと、前記入力帯域パワーと音声帯域パワーと雑音帯域パワーとに基き、音声信号の周波数帯域毎に重みをかけて背景雑音を抑圧する修正スペクトル減算ステップと、を具備することを特徴とする雑音抑圧方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は複数のマイクロホンを用いて雑音を抑圧し、目的の音声を取り出す雑音抑圧装置に関する。

【0002】

【従来の技術】環境下には種々の雑音源があることから、マイクロホンで音声信号を取り込む場合においても、周囲から紛れ込む雑音を避けることは難しい。しかし、雑音が混入した音声信号を再生すると、目的の音声は聴き辛いものとなるから、雑音成分の低減処理が必要となる。

【0003】ところで、音声に紛れる雑音の低減処理技術として、従来より知られているものに、複数のマイクロホンを用いて雑音を抑圧する技術がある。そして、このマイクロホン処理技術は、音声認識装置やテレビ会議装置などの音声入力を目的として従来から多くの研究者によって技術開発に力が注がれている。中でも、少ないマイクロホン数で大きな効果が得られる適応ビームフォーマ処理技術を利用したマイクロホンアレイに関しては、文献 1（電子情報通信学会編：音響システムとデジタル処理）あるいは文献 2（Heykin 著；Adaptive Filt

er Theory (Plentice Hall)) に述べられているように、一般化サイドロープキャンセラ (GSC)、フロスト型ビームフォーマ、参照信号法など、種々の方法が知られている。

【0004】なお、適応ビームフォーマ処理と云うのは、一般には、妨害雑音の到来方向に死角を形成したフィルタにより雑音を抑圧する処理である。しかしながら、この適応ビームフォーマ処理技術においては、実際の目的信号の到来方向が、仮定した到来方向と異なる場合、その目的信号が雑音と見做されて除去されてしまうことから、性能が劣化するという問題を抱えている。

【0005】そこで、これを改善すべく、例えば文献 3（宝珠山他：“ブロッキング行列にリーク適応フィルタを用いたロバスト一般化サイドロープキャンセラ”、電子情報通信学会論文誌 A Vol. J79-A No. 9 pp1516-1524 (1996. 9)）に開示されているように、仮定した到来方向と実際の到来方向とのずれを許容するような技術が開発されているが、この場合、目的信号の除去は軽減されても、実際の到来方向と仮定した到来方向とのずれにより、目的信号が歪むおそれがある。

【0006】これに対し、例えば、特開平 9-9794 号公報において、複数のビームフォーマを用いて、話者方向を逐次検知してその方向にビームフォーマの入力方向を修正することで、話者の方向を追尾し、目的信号の歪みを小さくする方法も開示されている。

【0007】しかしながら、特開平 9-9794 号公報に開示されている方法は、時間領域の適応フィルタ処理を行っているため、フィルタ係数から話者方向を推定する際、時間領域のフィルタ係数から周波数領域への変換が必要であり、計算量が大きくなる。

【0008】

【発明が解決しようとする課題】音声の雑音を抑圧する技術として、複数本のマイクロホンを用い、これらのマイクロホンで、話者の音声を取り込むと共に、妨害雑音の到来方向に死角を形成したフィルタを通すことにより、雑音成分を抑圧する適応ビームフォーマ処理技術がある。

【0009】この適応ビームフォーマ処理技術においては、実際の目的信号の到来方向、すなわち、話者のいる方向が、予め仮定した到来方向と異なる場合、目的信号が雑音と見做されて除去され、音声収集性能が劣化するという問題を抱えている。

【0010】そこで、これを改善すべく、仮定した到来方向と実際の到来方向とのずれを許容するような技術が開発されているが、この場合、目的信号の除去は軽減されても、実際の到来方向と仮定した到来方向とのずれにより、目的信号が歪む心配があり、得られる音声の品質の問題を残している。

【0011】また、複数のビームフォーマを用い、話者

方向を逐次検知してその方向にビームフォーマの入力方向を修正することで、話者の方向を追尾し、目的信号の歪みを小さくする方法も提案されている。しかしながら、この方法は、時間領域の適応フィルタ処理を行っているため、フィルタ係数から話者方向を推定する際、時間領域のフィルタ係数から周波数領域への変換が必要であり、計算量が大きくなるという問題があった。

【0012】故に、従来の技術はいずれも一長一短であり、高品位に目的信号を収集できると共に、処理時間も短時間で済むようなビームフォーマ処理技術の開発が囑望されている。

【0013】そこで、この発明の目的とするところは、周波数領域で動作するビームフォーマを用いることで、計算量を大幅に削減する雑音抑圧処理装置および雑音抑圧処理方法を提供することにある。

【0014】

【課題を解決するための手段】上記目的を達成するため、本発明は次のように構成する。

【0015】〔1〕 第1には、話者の発声した音声を少なくとも異なる2箇所以上の位置で受音する音声入力手段と、前記受音位置に対応する音声信号のチャンネル毎に周波数分析を行って複数チャンネルの周波数成分を出力する周波数分析手段と、この周波数分析手段にて得られる前記複数チャンネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声以外の音声を抑圧する到来雑音抑圧処理を行い、目的音声成分を得る第1のビームフォーマ処理手段と、前記周波数分析手段にて得られる前記複数チャンネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声を抑圧し、雑音成分を得る第2のビームフォーマ処理手段と、前記第1のビームフォーマ処理手段で計算されるフィルタ係数から雑音方向を推定する雑音方向推定手段と、前記第2のビームフォーマ処理手段で計算されるフィルタ係数から目的音方向を推定する目的音方向推定手段と、前記第1のビームフォーマにおいて入力対象となる目的音の到来方向である第1の入力方向を、前記目的音方向推定手段で推定された目的音方向に基づいて逐次修正する目的音方向修正手段と、前記第2のビームフォーマにおいて入力対象とする雑音の到来方向である第2の入力方向を、前記雑音方向推定手段で推定された雑音方向に基づいて逐次修正する雑音方向修正手段とを具備する。

【0016】〔2〕 また、第2には、本発明は、話者の発声した音声を少なくとも異なる2箇所以上の位置で受音する音声入力手段と、前記受音位置に対応する音声信号のチャンネル毎に周波数分析を行って複数チャンネルの周波数成分を出力する周波数分析手段と、この周波数分析手段にて得られる前記複数チャンネルの周波数成分につ

て、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声以外の音声を抑圧する到来雑音抑圧処理を行い、目的音声成分を得る第1のビームフォーマ処理手段と、前記周波数分析手段にて得られる前記複数チャンネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声を抑圧し、第1の雑音成分を得る第2のビームフォーマ処理手段と、前記周波数分析手段にて得られる前記複数チャンネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声を抑圧し、第2の雑音成分を得る第2のビームフォーマ処理手段と、前記第1のビームフォーマ処理手段で計算されるフィルタ係数から雑音方向を推定する雑音方向推定手段と、前記第2のビームフォーマ処理手段で計算されるフィルタ係数から第1の目的音方向を推定する第1の目的音方向推定手段と、前記第3の適応ビームフォーマ処理手段で計算されるフィルタ係数から第2の目的音方向を推定する第2の目的音方向推定手段と、前記第1のビームフォーマにおいて入力対象とする目的音の到来方向である第1の入力方向を、前記第1の目的音方向推定手段で推定された第1の目的音方向と、第2の目的音方向推定手段で推定された第2の目的音方向のいずれか一方または両方に基づいて逐次修正する第1の入力方向修正手段と、前記雑音方向修正手段で推定された雑音方向が所定の第1の範囲にある場合に、前記第2のビームフォーマにおいて入力対象とする雑音の到来方向である第2の入力方向を該雑音方向に基づいて逐次修正する第2の入力方向修正手段と、前記雑音方向修正手段で推定された雑音方向が所定の第2の範囲にある場合に、前記第3のビームフォーマにおいて入力対象とする雑音の到来方向である第3の入力方向を該雑音方向に基づいて逐次修正する第3の入力方向修正手段と、前記雑音方向推定手段で推定された雑音方向が所定の第1の範囲から到来したか所定の第2の範囲から到来したかに基づいて前記第1の出力雑音と前記第2の出力雑音のいずれか一方を真の雑音出力と決定していずれか一方の雑音を出力すると同時に、第1の音声方向推定手段と第2の音声方向推定手段のいずれの推定結果が有効であるかを決定していずれか一方の音声方向推定結果を第1の入力方向修正手段へ出力する有効雑音決定手段とを具備する。

【0017】〔3〕 更に第3には、本発明は、上記

〔1〕項または〔2〕項いずれかに記載の雑音抑圧装置において、前記得られた音声周波数を、周波数帯域毎に分割して帯域毎の音声パワーを計算する音声帯域パワー計算手段と、前記得られた雑音周波数成分を、周波数帯域毎に分割して帯域毎の雑音パワーを計算する雑音帯域パワー計算手段と、前記音声帯域パワー計算手段と雑音

帯域パワー計算手段とから得られる音声と雑音の周波数帯域パワーに基き、音声信号の周波数帯域毎に重みをかけて背景雑音を抑圧するスペクトル減算手段とからなるスペクトル減算雑音抑圧手段をさらに具備することを特徴とする。

【0018】 [4] 更に第4には、本発明は、上記

[1] 項または [2] 項いずれかに記載の雑音抑圧装置において、前記得られた音声周波数を、周波数帯域毎に分割して帯域毎の音声パワーを計算する音声帯域パワー計算手段と、前記得られた雑音周波数成分を、周波数帯域毎に分割して帯域毎の雑音パワーを計算する雑音帯域パワー計算手段と、前記音声入力手段から得られた入力信号を周波数分析した入力信号の周波数成分を周波数帯域毎に分割し、帯域毎の入力パワーを計算する入力帯域パワー計算手段と、前記入力帯域パワーと音声帯域パワーと雑音帯域パワーとに基き、音声信号の周波数帯域毎に重みをかけて背景雑音を抑圧する修正スペクトル減算手段を具備することを特徴とする。

【0019】そして、上記 [1] の構成の場合、話者の発声した音声を異なる2箇所以上の位置で音声入力手段は受音し、周波数分析手段では、これを前記受音位置に対応する音声信号のチャンネル毎に周波数分析して複数チャンネルの周波数成分を出力する。そして、第1のビームフォーマ処理手段はこの周波数分析手段にて得られる前記複数チャンネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声以外の音声を抑圧する到来雑音抑圧処理を行い、目的音声成分を得、また、第2のビームフォーマ処理手段は、前記周波数分析手段にて得られる前記複数チャンネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声を抑圧し、雑音成分を得る。そして、雑音方向推定手段は、前記第1のビームフォーマ処理手段で計算されるフィルタ係数から雑音方向を推定し、目的音方向推定手段は、前記第2のビームフォーマ処理手段で計算されるフィルタ係数から目的音方向を推定する。目的音方向修正手段は、前記第1のビームフォーマにおいて入力対象となる目的音の到来方向である第1の入力方向を、前記目的音方向推定手段で推定された目的音方向に基づいて逐次修正するので、第1のビームフォーマは第1の入力方向以外から到来する雑音成分を抑圧して話者の音声成分を低雑音で抽出することになる。また、雑音方向修正手段は、前記第2のビームフォーマにおいて入力対象とする雑音の到来方向である第2の入力方向を、前記雑音方向推定手段で推定された雑音方向に基づいて逐次修正するので、第2のビームフォーマは第2の入力方向以外から到来する成分を抑圧して話者の音声成分を抑圧した残りの雑音成分を抽出することになる。

【0020】このように本システムは雑音成分を抑圧した音声周波数成分と、音声成分を抑圧した雑音周波数成分とを別々に得ることができるが、この発明の最大の特徴は、第1及び第2のビームフォーマとして、周波数領域で動作するビームフォーマを用いるようにした点にある。そして、このことによって、計算量を大幅に削減することができるようにしている。

【0021】そしてこの発明によると、適応フィルタの処理量が大幅に低減されるのに加え、入力音声に対する周波数分析以外の周波数分析処理を省略することができ、かつ、フィルタ演算時に必要であった時間領域から周波数領域への変換処理も不要となり、全体の演算量を大幅に削減することができる。

【0022】すなわち、従来技術では、ビームフォーマで抑圧できない拡散性雑音の抑圧処理のために、スペクトルサブトラクション（以後、SSと略称する）処理を、ビームフォーマ処理の後に行うようにしており、このSSは周波数スペクトルを入力とするため、FFT（高速フーリエ変換）などの周波数分析が従来必要であったが、周波数領域で動作するビームフォーマを用いると当該ビームフォーマからは周波数スペクトルが出力されるため、これをSSに流用できるので、特別にSSのためのFFTを実施する従来のFFT処理工程は省略することができる。故に、全体の演算量を大幅に削減することができる。

【0023】また、ビームフォーマのフィルタを用いた方向推定の際に必要なであった時間領域から周波数領域への変換処理も不要となり、全体の演算量を大幅に削減することができる。

【0024】また、[2] の構成の場合、話者の発声した音声を異なる2箇所以上の位置で音声入力手段は受音し、周波数分析手段では、これを前記受音位置に対応する音声信号のチャンネル毎に周波数分析して複数チャンネルの周波数成分を出力する。そして、第1のビームフォーマ処理手段はこの周波数分析手段にて得られる前記複数チャンネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声以外の音声を抑圧する到来雑音抑圧処理を行い、目的音声成分を得、また、第2のビームフォーマ処理手段は、前記周波数分析手段にて得られる前記複数チャンネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声を抑圧し、雑音成分を得る。そして、雑音方向推定手段は、前記第1のビームフォーマ処理手段で計算されるフィルタ係数から雑音方向を推定し、目的音方向推定手段は、前記第2のビームフォーマ処理手段で計算されるフィルタ係数から目的音方向を推定する。

【0025】また、第1の目的音方向推定手段は前記第

2のビームフォーマ処理手段で計算されるフィルタ係数から第1の目的音方向を推定し、第2の目的音方向推定手段は、前記第3の適応ビームフォーマ処理手段で計算されるフィルタ係数から第2の目的音方向を推定する。

【0026】第1の入力方向修正手段は、前記第1のビームフォーマにおいて入力対象とする目的音の到来方向である第1の入力方向を、前記第1の目的音方向推定手段で推定された第1の目的音方向と、第2の目的音方向推定手段で推定された第2の目的音方向のいずれか一方または両方に基づいて逐次修正する。そして、第2の入力方向修正手段は、前記雑音方向修正手段で推定された雑音方向が所定の第1の範囲にある場合に、前記第2のビームフォーマにおいて入力対象とする雑音の到来方向である第2の入力方向を該雑音方向に基づいて逐次修正し、第3の入力方向修正手段は、前記雑音方向修正手段で推定された雑音方向が所定の第2の範囲にある場合に、前記第3のビームフォーマにおいて入力対象とする雑音の到来方向である第3の入力方向を該雑音方向に基づいて逐次修正する。従って、第2の入力方向修正手段の出力により第2の入力方向を修正される第2のビームフォーマは第2の入力方向以外から到来する成分を抑圧して残りの雑音成分を抽出することになり、また、第3の入力方向修正手段の出力により第3の入力方向を修正される第3のビームフォーマは第3の入力方向以外から到来する成分を抑圧して残りの雑音成分を抽出することになる。

【0027】そして、有効雑音決定手段は、前記雑音方向推定手段で推定された雑音方向が所定の第1の範囲から到来したか所定の第2の範囲から到来したかに基づいて前記第1の出力雑音と前記第2の出力雑音のいずれか一方を真の雑音出力と決定していずれか一方の雑音を出力すると同時に、第1の音声方向推定手段と第2の音声方向推定手段のいずれの推定結果が有効であるかを決定して有効な方の音声方向推定結果を第1の入力方向修正手段へ出力する。この結果、目的音方向修正手段は、前記第1のビームフォーマにおいて入力対象となる目的音の到来方向である第1の入力方向を、前記決定した方の目的音方向推定手段で得た目的音方向に基づいて逐次修正するので、第1のビームフォーマは第1の入力方向以外から到来する雑音成分を抑圧して話者の音声成分を低雑音で抽出することになる。

【0028】このように本システムは雑音成分を抑圧した音声周波数成分と、音声成分を抑圧した雑音周波数成分とを別々に得ることができるが、この発明の最大の特徴は、第1及び第2のビームフォーマとして、周波数領域で動作するビームフォーマを用いるようにした点にある。そして、このことによって、計算量を大幅に削減することができるようにしている。

【0029】そしてこの発明によると、適応フィルタの処理量が大幅に低減されるのに加え、入力音声に対する

周波数分析以外の周波数分析処理を省略することができ、かつ、フィルタ演算時に必要であった時間領域から周波数領域への変換処理も不要となり、全体の演算量を大幅に削減することができる。

【0030】また、本発明では、雑音追尾に監視領域を全く異ならせた雑音追尾用のビームフォーマを設けてあり、それぞれの出力からそれぞれ音声方向を推定させると共に、それぞれの推定結果からいずれが有効な雑音追尾をしているかを判断して、有効と判断された方のビームフォーマのフィルタ係数による音声方向の推定結果を第1の目的音方向修正手段に与えることで第1の目的音方向修正手段は、前記第1のビームフォーマにおいて入力対象となる目的音の到来方向である第1の入力方向を、前記目的音方向推定手段で推定された目的音方向に基づいて逐次修正するので、第1のビームフォーマは第1の入力方向以外から到来する雑音成分を抑圧して話者の音声成分を低雑音で抽出することができ、雑音源が移動してもこれを見失うことなく追尾して抑圧することができるようになるものである。

【0031】従来技術においては、2ch、すなわち、2本のマイクロホンだけでも目的音源の追尾を可能とすべく、雑音追尾用のビームフォーマを雑音抑圧のビームフォーマとは別に1個用いるが、例えば、雑音源が目的音の方向を横切って移動したような場合、雑音の追尾精度が低下することがあった。

【0032】しかし、本発明では、雑音を追尾するビームフォーマを複数用いて各々別個の追尾範囲を受け持つようにしたことにより、上記のような場合でも追尾精度の低下を抑止できるようになる。

【0033】また、[3]項の構成の場合、音声帯域パワー計算手段は、得られた音声周波数のスペクトル成分を、周波数帯域毎に分割して帯域毎の音声パワーを計算し、雑音帯域パワー計算手段は、前記得られた雑音周波数のスペクトル成分を、周波数帯域毎に分割して帯域毎の雑音パワーを計算する。そして、スペクトル減算手段は、前記音声帯域パワー計算手段と雑音帯域パワー計算手段とから得られる音声と雑音の周波数帯域パワーに基づき、音声信号の周波数帯域毎に重みをかけて背景雑音を抑圧する。

【0034】この構成によれば、ビームフォーマでは抑圧できない方向性のない雑音（背景雑音）は、本発明システムのビームフォーマで得ることのできる目的音声成分と雑音成分を利用し、これをスペクトルサブトラクション処理することで抑圧する。すなわち、本システムでは、ビームフォーマとして目的音声成分抽出用と雑音成分抽出用の2つのビームフォーマを備えているが、これらのビームフォーマの出力である目的音声成分と雑音成分を利用してスペクトルサブトラクション処理することにより、方向性のない背景雑音成分の抑圧を行う。スペクトルサブトラクション（SS）処理は雑音抑圧処理と

して知られるが、一般的に行われるスペクトルサブトラクション (SS) 処理は、1 チャンネルのマイクロホン (つまり、1 本のマイクロホン) を用い、このマイクロホンの出力から音声のない区間において雑音のパワーを推定するため、非定常な雑音が音声に重畳している場合には対処できない。また、2 チャンネルのマイクロホン (つまり、2 本のマイクロホン) を用いて、一方を雑音収集用、片方を雑音重畳音声収集用とする場合にも、両マイクロホンの設置場所を離す必要があり、その結果、音声に重畳する雑音と、雑音収集用マイクロホンで取り込む雑音との位相がずれ、スペクトルサブトラクション処理しても雑音抑圧の改善効果は大きく上がらない。

【0035】しかし、本発明では、雑音成分を取り出すビームフォーマを用意して、このビームフォーマの出力を用いるようにしたため、位相のずれは補正されており、従って、非定常雑音の場合でも高精度なスペクトルサブトラクション処理を実現できる。さらに、周波数領域のビームフォーマの出力を利用しているため、周波数分析を省略してスペクトルサブトラクションが可能であり、従来より少ない演算量で非定常雑音を抑圧できる。

【0036】更に [4] 項の発明は、上記 [3] の発明の雑音抑圧装置において、音声入力手段から得られた入力信号を周波数分析した入力信号の周波数成分を周波数帯域毎に分割し、帯域毎の入力パワーを計算する入力帯域パワー計算手段を設けて、スペクトル減算手段には、入力帯域パワーと音声帯域パワーと雑音帯域パワーとに基き、音声信号の周波数帯域毎に重みをかけて背景雑音を抑圧する処理を実施させるようにしており、この構成の場合、音声帯域パワー計算手段は、得られた音声周波数のスペクトル成分を、周波数帯域毎に分割して帯域毎の音声パワーを計算し、雑音帯域パワー計算手段は、前記得られた雑音周波数のスペクトル成分を、周波数帯域毎に分割して帯域毎の雑音パワーを計算する。また、入力帯域パワー計算手段があり、この入力帯域パワー計算手段は、音声入力手段から得られた入力信号を周波数分析して得た入力音声の周波数スペクトル成分を受けて、これを周波数帯域毎に分割し、帯域毎の入力パワーを計算する。そして、スペクトル減算手段は、前記音声帯域パワー計算手段と雑音帯域パワー計算手段とから得られる音声と雑音の周波数帯域パワーに基き、音声信号の周波数帯域毎に重みをかけて背景雑音を抑圧する。

【0037】この [4] 項の発明においては、[3] 項の発明におけるスペクトルサブトラクション (SS) 処理において、更に雑音成分についてそのパワーを修正するようにしたことにより、一層高精度に雑音抑圧を行うことを可能とするものである。すなわち、[3] 項の発明では雑音源のパワー N が小さいという仮定をおいたため、スペクトルサブトラクション (SS) 処理を行うと雑音源の成分が音声に重畳している部分では歪みが大きくなる可能性が残るが、ここでは入力信号のパワーを用

いて第 3 の発明でのスペクトルサブトラクション処理における帯域重みの計算を修正するようにした。これにより、方向を持つ雑音成分および方向のない雑音成分を抑圧した歪みの少ない音声成分のみの抽出ができるようになる。

【0038】

【発明の実施の形態】以下、本発明の実施の形態につき図面を参照して説明する。

【0039】(実施例 1) はじめに、実施例 1 について説明する。この実施例 1 は請求項 1 の内容に相当する。

【0040】図 1 は実施例 1 のシステムの構成例を示すブロック図であって、本発明の一実施形態に係る雑音抑圧装置の基本構成を示すブロック図である。本発明は、マイクロホン数が $2ch$ (ch ; チャンネル) すなわち、2 本と云った最小の場合でも話者追尾可能とするための技術であるため、ここでは $2ch$ で説明するが、 $3ch$ 以上となった場合でも処理の方法は同様である。

【0041】図 1 において、11 は音声入力部、12 は周波数解析部、13 は第 1 のビームフォーマ、14 は第 1 の入力方向修正部、15 は第 2 の入力方向修正部、16 は第 2 のビームフォーマ、17 は雑音方向推定部、18 は目的音方向推定部 (音声方向推定部) である。

【0042】これらのうち、音声入力部 11 は、例えば、音声収集対象である話者の発声した音声 (目的音声) を異なる 2 箇所以上の位置で受音するためのものであり、具体的にはそれぞれ地点を異ならせて設置した 2 本のマイクロホンを用いて音声を取り込み、電気信号に変換するものである。また、周波数解析部 12 は、前記マイクロホンの受音位置に対応する音声信号のチャンネル毎に周波数分析を行って複数チャンネルの周波数成分を出力するものであり、具体的にはここでは第 1 のマイクロホンのとらえた音声信号 (第 1 チャンネル $1ch$ の音声信号) および第 2 のマイクロホンのとらえた音声信号 (第 2 チャンネル $2ch$ の音声信号) を、それぞれ別々に高速フーリエ変換するなどして時間領域の信号成分から周波数領域の成分のデータに変換することにより、各チャンネル別に周波数スペクトルのデータに変換して出力するものである。

【0043】第 1 のビームフォーマ 13 は、この周波数解析部 12 からの複数チャンネルの周波数成分出力、この場合、 $1ch$ 、 $2ch$ の音声信号を用いて、これより目的音声の周波数成分を抽出するためのものであって、前記 $1ch$ 、 $2ch$ それぞれの周波数成分 (周波数スペクトルデータ) を用いて適応フィルタ処理により目的の音声以外の到来雑音の抑圧処理を行うことにより、目的とする音源方向からの周波数成分を抽出するといったことを行う処理手段であり、第 2 のビームフォーマ 16 は、周波数解析部 12 からの複数チャンネルの周波数成分出力、この場合、 $1ch$ 、 $2ch$ の音声信号を用いて、これより雑音源方向からの周波数成分を抽出するためのも

のであって、前記1ch, 2chそれぞれの周波数成分（周波数スペクトルデータ）を用いて適応フィルタ処理により雑音音源方向からの音声以外の成分の抑圧処理を行うことにより、雑音源方向からの周波数スペクトル成分のデータを抽出するといったことを行う処理手段である。

【0044】また、雑音方向推定部17は、前記第1のビームフォーマ13で計算されるフィルタ係数から雑音方向を推定すると云った処理を行うものであって、具体的には前記第1のビームフォーマ13の適応フィルタから得られるフィルタリング処理用のフィルタ係数などのパラメータを用いて雑音方向を推定し、その推定量対応のデータを出力し、また、目的音方向推定部（音声方向推定部）18は、前記第2のビームフォーマ16で計算されるフィルタ係数から目的音方向を推定すると云った処理を行うものであって、具体的には前記第2のビームフォーマ16の適応フィルタで用いられているフィルタ係数などのパラメータから雑音方向を推定し、その推定量対応のデータを出力するものである。

【0045】また、第1の入力方向修正部14は、本来の目的音方向にビームフォーマの入力方向を修正するためのものであって、前記第1のビームフォーマ13において、入力対象とする目的音の到来方向である第1の入力方向を、前記目的音方向推定部18で推定された目的音方向に基づいて逐次方向修正するための出力を発生し、第1のビームフォーマ13に与えるものである。具体的には、第1の入力方向修正部14は、目的音方向推定部18の出力する推定量対応のデータを現在の目的とする音源方向の角度情報 α に変換して目標角度情報 α として第1のビームフォーマ13に出力するものである。

【0046】第2の入力方向修正部15は第2のビームフォーマ16の入力方向を雑音方向に修正するためのものであって、前記第2のビームフォーマ16において、入力対象とする雑音の到来方向である第2の入力方向を、前記雑音方向推定部17で推定された雑音方向に基づいて逐次方向修正するための出力を発生し、第2のビームフォーマ14に与えるものである。具体的には、第2の入力方向修正部15は、雑音方向推定部17の出力する推定量対応のデータを現在の目的とする雑音源方向の角度情報に変換して目標角度情報 α として第2のビームフォーマ16に出力するものである。

【0047】ここでビームフォーマ13, 16の構成例を示しておく。

<ビームフォーマの構成例>本発明システムで用いるビームフォーマ13, 16は、図2(a)に示すような構成となる。すなわち、本発明システムにおいて用いられるビームフォーマ13, 16は、入力音声から抽出したい対象となる信号成分を得ることができるようにするために、抽出したい対象となる信号成分の到来方向に、ビームフォーマの入力方向を設定するための移相部10

0と、抽出したい対象となる信号成分の到来方向以外の方向からの成分を抑圧するビームフォーマ本体101とから構成される。

【0048】移相部100は補正ベクトル生成部100aと乗算手段100b, 100cとから構成され、ビームフォーマ本体101は加算手段101a, 101b, 101cと適応フィルタ101dとから構成される。

【0049】補正ベクトル生成部100aは入力方向修正部14または15からの角度情報 α を入力方向の情報として受けて、これより α 対応の補正ベクトルを生成するものであり、乗算手段100bは周波数分析部12から出力されるch1の周波数スペクトル成分のデータに対して補正ベクトル分を乗算して出力するものであり、乗算手段100cは周波数分析部12から出力されるch2の周波数スペクトル成分のデータに対して補正ベクトル分を乗算して出力するものである。

【0050】また、加算手段101aは乗算手段100bの出力と加算手段100cの出力を加算して出力するものであり、加算手段101bは乗算手段100bの出力と加算手段100cの出力の差分を出力するものであり、加算手段101cは加算手段101aの出力に対する適応フィルタ101dの出力の差分をビームフォーマの出力として出力するものであり、適応フィルタ101dは加算手段101bの出力に対してフィルタリング演算処理して出力するためのデジタルフィルタであって、加算手段101cの出力が最小となるようにフィルタ係数（パラメータ）が逐次変更される構成である。

【0051】ここで、本例ではマイクロホン構成が2本、すなわち、第1及び第2のマイクロホンm1, m2を用いる収集音声2チャンネル(ch1, ch2)構成のシステムとしており、この場合、ビームフォーマの入力方向の設定とは、図2(b)に示すように、入力対象の存在する方向からの音声信号が等価的に同時に両マイクロホンm1, m2に到着したと見做せるように、ch1, ch2の2つの音声チャンネルの周波数成分に対して遅延を施し、位相を揃える（整相）ようにすることを指す。これは、図2の構成の場合、入力方向修正部14, 15の出力する角度情報 α 対応に移相部100で移相調整することによって実現している。

【0052】すなわち、図2の構成の場合、移相部100は補正したい入力方向（角度情報 α ）対応の補正ベクトルを補正ベクトル生成部100aで生成するようにしており、この補正ベクトルを1ch, 2chの各チャンネルの信号にそれぞれ乗算する乗算手段100b, 100cで乗算する構成とした移相部100により次のようにして位相を揃える。

【0053】例えば、図2(b)に符号m1, m2を付して示すような無指向性マイクロホン配置であって、P1点に居る目的音源である話者が、あたかもP2点に居るかのよう信号に位相補正することを考えてみる。こ

のような場合には、距離 d だけ離れた第1のマイクロホン $m1$ で検出した話者音声信号($ch1$)の位相と第2のマイクロホン $m2$ で検出した話者音声信号($ch2$)の位相が同じになるように、第1のマイクロホン $m1$ の話者音声信号($ch1$)に伝搬時間差 τ

$$\tau = r \cdot c = r \cdot \sin \alpha$$

$$r = d \cdot \sin \alpha$$

に相当する複素数 $W1$

$$W1 = (\cos j\omega\tau, \sin j\omega\tau)$$

の複素共役をかける。ここで、 c は音速、 d はマイクロホン間距離、 α はマイクロホン $m1$ から見た目的音の音源である話者の移動した角度、 j は虚数、 ω は角周波数である。

【0054】つまり、 $W1$ の複素共役をかけたことにより、 α なる角度に移動した目的音源の音声について注目すれば、第1のマイクロホン $m1$ でとらえた信号($ch1$)が、第2のマイクロホン $m2$ でとらえた信号と同位相となるように移相制御したことになる。

【0055】尚、第2のマイクロホン $m2$ の信号($ch2$)には、複素数 $W2 = (1, 0)$ の複素共役をかけるものとする。つまり、これは第2のマイクロホン $m2$ の信号($ch2$)には、角度補正をしないことを意味する。

【0056】ここで、複素数 $W1$ と複素数 $W2$ を並べたベクトル $\{W1, W2\}$ は、一般に方向ベクトルと呼ばれ、この $\{W1, W2\}$ における複素共役のベクトル共役 $\{W1^*, W2^*\}$ を、補正ベクトルと呼ぶ。

【0057】角度情報 α 対応に補正ベクトルを生成させ、 $ch1$ 、 $ch2$ の周波数スペクトル成分に対してこの補正ベクトルを乗算すれば、第1のマイクロホン $m1$ の出力は、音源が $P1$ より $P2$ に移動したにもかかわらず、第2のマイクロホン $m2$ の位相と同じになるように補正されたことになり、第1のマイクロホン $m1$ に関する限り、第2のマイクロホン $m1$ 、 $m2$ の $P2$ 位置音源に対する距離はあたかも等しいかたちになる。

【0058】本実施例では、ビームフォーマは2つあるが、これら2つあるビームフォーマのうち、第1のビームフォーマ13はその移相部100により目的音の音源方向を入力対象方向とするように、 $ch1$ (もしくは $ch2$)の周波数成分に上述の手法で遅延を施し、第2のビームフォーマ16はその移相部100により雑音源方向を入力対象方向とするように、 $ch1$ (もしくは $ch2$)の周波数成分に上述の手法で遅延を施してそれぞれ両者の位相を揃える。ただし、目的音 S の到来方向以外からの音成分、すなわち、雑音成分 N については第1および第2のマイクロホン $m1$ 、 $m2$ ともに位相は全く無修正であるから、第1のマイクロホン $m1$ と第2のマイクロホン $m2$ で検出されるタイミングに時間差がある。

【0059】このように移相部100により、目的音方向の音源からの検出される音声信号について位相修正し

た第1のマイクロホン $m1$ の出力(目的音声成分 S と雑音成分 N からなる $ch1$ の周波数スペクトルデータ)および修正の加えられない第2のマイクロホン $m2$ の出力

(目的音声成分 S と雑音成分 N' からなる $ch2$ の周波数スペクトルデータ)は、それぞれ加算手段101a、101bに入力される。そして、加算手段101aでは $ch1$ の出力と $ch2$ の出力が加算されることによって目的音声 S の2倍の信号と雑音成分 $N+N'$ についてのパワー成分が求められ、加算手段101bでは $ch1$ の出力($S+N$)と $ch2$ の出力($S+N'$)の差分($(S+N) - (S+N') = N-N'$)、つまり、ノイズ分のパワー成分が求められる。そして、加算手段101cで加算手段101aの出力に対する適応フィルタ101dの出力の差分を求め、これをビームフォーマの出力とすると共に、適応フィルタ101dにフィードバックする。

【0060】適応フィルタ101dは加算手段101bの出力に対して現在の探査方向対応の方向から到来した音の成分の周波数スペクトルが抽出されるようフィルタリング演算処理して出力するためのデジタルフィルタであり、逐次、角度 1° 刻みに到来信号の探査角度を可変して、入力される信号方向に探査角度が一致したとき最大の出力を出す。従って、到来信号の入射方向と探査角度が一致すれば適応フィルタ101dの出力($N-N'$)は最大になる。そして、適応フィルタ101dの出力($N-N'$)は雑音成分のパワーであるから、それが最大のときの出力を加算手段101cに与え、加算手段101aからの出力($2S+N+N'$)から差し引けば、雑音成分 N が最大限キャンセルされて雑音抑圧が成される。故に、この状態のときは、加算手段101cの出力は最小である。

【0061】そのため、適応フィルタ101dは加算手段101cの出力が最小となるように角度 1° 刻みの信号到来方向探査角度(角度 1° 刻みの方向別感度)とフィルタ係数(パラメータ)を逐次変更させることにより、到来信号の入射方向と探査角度(到来信号の入射方向とその方向に対する感度)が一致することになるから、適応フィルタ101dはこれらを制御しつつ、加算手段101cの出力が最小となるようにする。

【0062】つまり、この制御の結果、目的方向からの音声成分をビームフォーマは抽出できることになる。また、雑音成分を目的音として抽出する場合には、上述の目的音を雑音と見做すようにしたかたちで、上記制御を施すようにすればよい。

【0063】なお、ビームフォーマ本体101に関しては、一般化サイドローブキャンセラ(GSC)の他に、フロスト型ビームフォーマなど種々のものが上述同様の考え方で適用可能であり、従って、本発明では特に限定はされない。

【0064】このような構成の本システムの作用を説明

する。本システムは、目的音の音声周波数成分と雑音周波数成分とを別々に抽出出力する構成としていることを特徴としている。

【0065】まず、複数のマイクロホンを持つ音声入力部 11、この例では第 1 及び第 2 の計 2 本のマイクロホン m1、m2 を持つ音声入力部 11 で ch1、ch2 の音声を取り込む。そして、この音声入力部 11 から入力された 2 チャンネル分の音声の信号 ch1、ch2（すなわち、第 1 チャンネル ch1 は第 1 のマイクロホン m1 からの音声、第 2 チャンネル ch2 は第 2 のマイクロホン m2 からの音声に該当する）は、周波数分析部 12 に送られ、ここで例えば高速フーリエ変換（FFT）等の処理を行うことによって、それぞれのチャンネル別に周波数成分（周波数スペクトル）が求められる。

【0066】周波数分析部 12 でそれぞれ求められたチャンネル別の周波数成分は、それぞれ第 1 及び第 2 のビームフォーマ 13、16 に与えられる。

【0067】第 1 のビームフォーマ 13 では、2 チャンネル分の周波数成分入力について、目的音の方向対応に位相を合わせた上で、周波数領域の適応フィルタにより上述のようにして処理することで雑音を抑圧し、目的音の方向の周波数成分を出力する。

【0068】ここで、具体的に説明すると第 1 の入力方向修正部 14 は第 1 のビームフォーマ 13 に対して次のような角度情報（ α ）を与える。つまり、第 1 の入力方向修正部 14 は、与えられる音声方向推定部 18 からの出力を用い、目的音の方向があたかもマイクロホンの正面方向となるよう、上記 2 チャンネルの周波数成分の入力位相を整えるに必要な角度情報（ α ）を入力方向修正量として第 1 のビームフォーマ 13 に対して与える。

【0069】この結果、第 1 のビームフォーマ 13 はこの修正量（ α ）対応に目的音方向を修正し、当該目的音方向以外の方向から到来する音声を抑圧させるようにすることで、雑音成分を抑圧し、目的音を抽出する。

【0070】すなわち、目的音方向推定部 18 は雑音成分を抽出するための第 2 のビームフォーマ 16 における適応フィルタのパラメータを用いて雑音源方向を知り、それを反映させた出力を出し、第 1 の入力方向修正部 14 ではこの目的音方向推定部 18 からの出力対応に入力方向修正量（ α ）を発生してこの修正量（ α ）対応に第 1 のビームフォーマ 13 における目的音方向を修正し、これによって第 1 のビームフォーマ 13 に当該目的音方向以外の方向から到来する音声を抑圧させるようにすることで、雑音成分を抑圧し、目的音を抽出する。

【0071】つまり、第 2 のビームフォーマ 16 の場合、雑音が目的音であるから、雑音に位相を合わせている。その結果、第 2 のビームフォーマ 16 では話者の音源は雑音源として扱われ、ビームフォーマの内蔵する適応フィルタは話者音源からの音を抽出する処理をすることになるので、当該第 2 のビームフォーマ 16 の適応フ

ィルタのパラメータからは話者音源の方向を反映した出力が得られる。従って、目的音方向推定部 18 により、第 2 のビームフォーマ 16 における適応フィルタのパラメータを用いて雑音源方向を知れば、それは目的音である話者音源の方向を反映させたものである。従って、目的音方向推定部 18 により、第 2 のビームフォーマ 16 における適応フィルタのパラメータを反映させた出力を出し、第 1 の入力方向修正部 14 でこの目的音方向推定部 18 からの出力対応に入力方向修正量（ α ）を発生し、この修正量対応に第 1 のビームフォーマ 13 における目的音方向を修正すれば、第 1 のビームフォーマ 13 に当該目的音方向以外の方向から到来する音声を抑圧させることができる。

【0072】また、第 2 のビームフォーマ 16 では、2 チャンネル分の周波数成分入力に対して、周波数領域の適応フィルタにより目的音を抑圧し、雑音の方向の周波数成分を出力する。ここでは、具体的には雑音の方向をマイクロホンの正面と仮定し、2 つのマイクロホンに対して雑音が同時に到着したと見做せるように、雑音方向推定部 17 からの出力を用いて第 2 の入力方向修正部 5 で位相を整える操作（整相）を行う。

【0073】すなわち、雑音方向推定部 17 では、話者音声成分を抽出するための第 1 のビームフォーマ 13 における適応フィルタのパラメータを用いて雑音音源方向を知り、それを反映させた出力を出し、第 2 の入力方向修正部 15 では雑音方向推定部 17 からの出力対応に入力方向修正量（ α ）を発生させて第 2 のビームフォーマ 16 に与えることによって、当該第 2 のビームフォーマ 16 に当該修正量対応に雑音方向を修正させるようにし、この方向以外の方向から到来する音声を抑圧することで雑音成分のみを抽出する。

【0074】ここで、雑音方向推定部 17 では、第 1 のビームフォーマ 13 の適応フィルタから雑音方向を推定し、目的音方向推定部 18 では、第 2 のビームフォーマ 16 の適応フィルタから目的音方向を推定する。なお、これらの処理は、例えば、8 [msec] 等の短い固定時間毎に行われる。以降固定時間をフレームと呼ぶ。

【0075】このようにして、第 1 のビームフォーマ 13 により、目的音（話者）の音声成分を抽出することができ、また、第 2 のビームフォーマ 16 により、雑音成分を抽出することができる。

【0076】本装置の設置環境が、静かな会議室であり、この会議室にテレビ会議システム設置して当該テレビ会議システムの話者音声抽出のために使用しているとすれば、除去しなければならない雑音と云っても、そう問題のある大きな妨害音ではないと考えられるので、このような場合、第 1 のビームフォーマ 13 により、抽出された目的音（話者）の成分を逆フーリエ変換して時間領域に戻すことで音声信号に戻し、これをスピーカなどで音声として出力させたり、送信するなどすれ

ば、低雑音化された話者音声として利用できる。

【0077】ここで、方向推定部17、18の処理手順について触れておく。

【0078】＜方向推定部の処理手順＞図3に方向推定部17、18の処理手順を示す。

【0079】この処理はフレーム毎に行われる。まず、初期設定をする（ステップS1）。この初期設定内容としては図3に点線枠で囲んで示してあるように、“目的音の追尾範囲”を“ $0^\circ \pm \theta_r$ （例えば、 20° ）”とし、それ以外の範囲を雑音の探索範囲として設定する。

【0080】初期設定が終わったならば、次にステップS2の処理に移る。このステップS2では方向ベクトルを生成する処理を行う。そして、方向別感度計算を行った後、方向別感度周波数累積を行う（ステップS3、S4）。

【0081】そして、これを全ての周波数と方向について、実施した後、最小値であるものを求めて、その最小値となった累積値を持つものの方向を信号到来方向とする（ステップS5、S6）。

【0082】すなわち、具体的にはステップS2からS4においては、フィルタ係数 $W(k)$ と方向ベクトル $S(k, \theta)$ との内積を各周波数成分毎に 1° 刻みで所定の範囲の方向について計算し、対応する方向への感度を求め、次に、全周波数成分についてその感度を加算すると云う処理を行う。そして、ステップS7、S8においては、全周波数成分についてその感度を加算した結果として得られる各方向別の累積値のうち、その値が最小値である方向を、信号到来方向とすると云う処理をする。この図3に示した処理手順は、雑音方向推定部17および目的音推定部18ともに同様のものとなる。

【0083】このようにして、雑音方向推定部17は雑音方向の推定を行い、また、目的音推定部18は目的音方向の推定を行う。そして、この推定結果はそれぞれの対応する入力方向修正部14、15に与える。

【0084】雑音方向の推定結果を受け取った第1の入力方向修正部14は、前フレームまでの入力方向と現フレームの方向推定結果を平均化し、新たな入力方向を計算してビームフォーマの移相部100へ出力し、また、目的音推定結果を受け取った第2の入力方向修正部15もまた、前フレームまでの入力方向と現フレームの方向推定結果を平均化し、新たな入力方向を計算してビームフォーマの移相部100へ出力する。

【0085】平均化は例えば、係数 β を用いて次式のように行う。

$$\theta_1(n) = \theta_1(n-1) \cdot (1-\alpha) + E(n) \cdot \beta$$

ここで、 θ_1 は音の入力方向、 n は処理フレームの番号、 E は現フレームの方向推定結果である。なお、係数 β はビームフォーマの出力パワーに基づいて可変にしてもよい。

【0087】ビームフォーマがGSCである場合に、従来、方向推定の際、時間領域のフィルタ係数から周波数領域への変換が必要であったが、本発明ではGSCの適応フィルタが周波数スペクトルに対して方向性感度を以てフィルタ演算処理し、目的方向外の成分を抽出すると云った処理をするものを用いており、フィルタ演算処理に使用するフィルタ係数は、もともと周波数領域で得られるため、従来のように、時間領域のフィルタ係数から周波数領域への変換と云う処理は不要となる。従って、本発明システムではGSCは使用していても、時間領域のフィルタ係数から周波数領域への変換が不要である分、処理の高速化が可能となる。

【0088】＜全体の処理手順＞図4に実施例1に係るシステムの全体の処理手順を示す。この処理はフレーム毎に行われる。

【0089】まず、初期設定する（ステップS11）。初期設定内容としては、目的音方向の追尾範囲を $0^\circ \pm \theta_r$ （例えば $\theta_r = 20^\circ$ ）とし、雑音方向推定部の探索範囲を

$$\theta_r < \phi_1 < 180^\circ - \theta_r, \\ -180^\circ + \theta_r < \phi_1 < -\theta_r$$

とし、目的音方向推定部18の探索範囲を

$$-\theta_r < \phi_2 < \theta_r$$

とする。

【0090】そして、目的音の入力方向の初期値を $\theta_1 = 0^\circ$ 、雑音の入力方向の初期値を $\theta_2 = 90^\circ$ とする。

【0091】初期設定が済んだならば、まず、第1のビームフォーマ13の処理を行い（ステップS12）、雑音方向を推定し（ステップS13）、雑音方向が ϕ_2 の範囲内であれば、第2のビームフォーマ16の入力方向を修正し（ステップS14、S15）、そうでなければ修正しない（ステップS14）。

【0092】次に、第2のビームフォーマ16の処理に進み（ステップS16）、目的音の方向を推定する（ステップS17）。そして、この推定した目的音の方向が ϕ_1 の範囲内ならば、第1のビームフォーマ13の入力方向を修正し（ステップS18、S19）、そうでなければ何もせずに、次のフレームの処理に移る。

【0093】以上、実施例1においては、ビームフォーマとして周波数領域で動作するビームフォーマを用いるようにしたことを特徴としており、これによって計算量を大幅に削減することができるようにしたことを特徴としている。

【0094】すなわち、話者の発声した音声を少なくとも異なる2箇所以上の位置で受信する音声入力手段と、前記受信位置に対応する音声信号のチャネル毎に周波数分析を行って複数チャネルの周波数成分を出力する周波数分析手段と、この周波数分析手段にて得られる前記複数チャネルの周波数成分について、所望方向外の感度が

低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声以外の音声を抑圧する到来雑音抑圧処理を行い、目的音声成分を得る第1のビームフォーマ処理手段と、前記周波数分析手段にて得られる前記複数チャネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声を抑圧し、雑音成分を得る第2のビームフォーマ処理手段と、前記第1のビームフォーマ処理手段で計算されるフィルタ係数から雑音方向を推定する雑音方向推定手段と、前記第2のビームフォーマ処理手段で計算されるフィルタ係数から目的音方向を推定する目的音方向推定手段と、前記第1のビームフォーマにおいて入力対象となる目的音の到来方向である第1の入力方向を、前記目的音方向推定手段で推定された目的音方向に基づいて逐次修正する目的音方向修正手段と、前記第2のビームフォーマにおいて入力対象とする雑音の到来方向である第2の入力方向を、前記雑音方向推定手段で推定された雑音方向に基づいて逐次修正する雑音方向修正手段とを具備する。

【0095】そして、話者の発声した音声を異なる2箇所以上の位置で音声入力手段は受音し、周波数分析手段では、これを前記受音位置に対応する音声信号のチャネル毎に周波数分析して複数チャネルの周波数成分を出力する。そして、第1のビームフォーマ処理手段はこの周波数分析手段にて得られる前記複数チャネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声以外の音声を抑圧する到来雑音抑圧処理を行い、目的音声成分を得、また、第2のビームフォーマ処理手段は、前記周波数分析手段にて得られる前記複数チャネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声を抑圧し、雑音成分を得る。そして、雑音方向推定手段は、前記第1のビームフォーマ処理手段で計算されるフィルタ係数から雑音方向を推定し、目的音方向推定手段は、前記第2のビームフォーマ処理手段で計算されるフィルタ係数から目的音方向を推定する。目的音方向修正手段は、前記第1のビームフォーマにおいて入力対象となる目的音の到来方向である第1の入力方向を、前記目的音方向推定手段で推定された目的音方向に基づいて逐次修正するので、第1のビームフォーマは第1の入力方向以外から到来する雑音成分を抑圧して話者の音声成分を低雑音で抽出することになる。また、雑音方向修正手段は、前記第2のビームフォーマにおいて入力対象とする雑音の到来方向である第2の入力方向を、前記雑音方向推定手段で推定された雑音方向に基づいて逐次修正するので、第2のビームフォーマは第2の入力方向以外から到来する成分を抑圧して話者の音声成

分を抑圧した残りの雑音成分を抽出することになる。

【0096】このように本システムは雑音成分を抑圧した音声周波数成分と、音声成分を抑圧した雑音周波数成分とを別々に得ることができるが、この発明の最大の特徴は、第1及び第2のビームフォーマとして、周波数領域で動作するビームフォーマを用いるようにした点にある。そして、このことによって、計算量を大幅に削減することができるようにしている。

【0097】そしてこの発明によると、適応フィルタの処理量が大幅に低減されるのに加え、入力音声に対する周波数分析以外の周波数分析処理を省略することができ、かつ、フィルタ演算時に必要であった時間領域から周波数領域への変換処理も不要となり、全体の演算量を大幅に削減することができる。

【0098】すなわち、従来技術では、ビームフォーマで抑圧できない拡散性雑音の抑圧処理のために、スペクトルサブトラクション（以後、SSと略称する）処理を、ビームフォーマ処理の後に行うようにしており、このSSは周波数スペクトルを入力とするため、FFT（高速フーリエ変換）などの周波数分析が従来必要であったが、周波数領域で動作するビームフォーマを用いると当該ビームフォーマからは周波数スペクトルが出力されるため、これをSSに流用できるので、特別にSSのためのFFTを実施する従来のFFT処理工程は省略することができる。故に、全体の演算量を大幅に削減することができる。

【0099】また、ビームフォーマのフィルタを用いた方向推定の際に必要であった時間領域から周波数領域への変換処理も不要となり、全体の演算量を大幅に削減することができる。

【0100】次に、雑音源が目的音方向の範囲を横切って移動した場合にも追尾が高精度で行えるようにした例を実施例2として説明する。

【0101】（実施例2）本発明に係る第2の実施例について説明する。これは、請求項2の発明に相当する。

【0102】本例では、雑音源が目的音方向の範囲を横切って移動した場合にも追尾が高精度で行えるように、雑音を追尾するビームフォーマを2つ用いる場合の例について説明する。全体構成を図4に示す。図4において、11は音声入力部、12は周波数解析部、13は第1のビームフォーマ、14は第1の入力方向修正部、15は第2の入力方向修正部、16は第2のビームフォーマ、17は雑音方向推定部、18は第1の音声方向推定部（目的音方向推定部）、そして、21は第3の入力方向修正部、22は第3のビームフォーマ、23は第2の音声方向推定部、24は有効雑音決定部である。

【0103】これらのうち、第3の入力方向修正部21は、第3のビームフォーマ22の入力方向を雑音方向に修正するためのものであって、第3のビームフォーマ22において、入力対象とする雑音の到来方向である第3

の入力方向を、前記雑音方向推定部 17 で推定された雑音方向に基づいて逐次方向修正するための出力を発生し、第 3 のビームフォーマ 22 に与えるものである。具体的には、第 3 の入力方向修正部 21 は、雑音方向推定部 17 の出力する推定量対応のデータを現在の目的とする雑音源方向の角度情報に変換して目標角度情報 α として第 3 のビームフォーマ 22 に出力するものである。

【0104】第 3 のビームフォーマ 22 は、周波数分析部 12 からの複数チャンネルの周波数成分出力、この場合、1ch、2ch の音声信号の周波数スペクトルを用いて、これより雑音源方向からの周波数スペクトル成分を抽出するためのものであって、前記 1ch、2ch それぞれの周波数成分（周波数スペクトルデータ）に対して方向別感度調整を施した適応フィルタ処理により雑音音源方向以外の周波数スペクトル成分の抑圧処理を行うことで、雑音音源方向からの周波数スペクトル成分のデータを抽出するといったことを行う処理手段である。この第 3 のビームフォーマ 22 も第 1 及び第 2 のビームフォーマ 13、16 同様、図 2 で説明した如きの構成を採用している。

【0105】第 2 の音声方向推定部 23 は、目的音声推定部（音声方向推定部）18 と同様のものであって、前記第 3 のビームフォーマ 22 で計算されるフィルタ係数から目的音方向を推定すると云った処理を行うものであり、具体的には前記第 3 のビームフォーマ 22 の適応フィルタから音声方向を推定し、その推定量対応のデータを出力するものである。

【0106】有効雑音決定部 24 は、音声方向推定部 18、23 および雑音方向推定部 17 の推定する音声方向および雑音方向の情報に基づき、第 2 のビームフォーマ 16 と第 3 のビームフォーマ 22 のいずれが雑音を有効に追尾しているかを判断し、有効に追尾していると判断した方のビームフォーマの出力を、雑音成分として出力するものである。なお、その他、図 1 の構成と同一符号を付したものは同一物を示しているので、詳細は先の説明を参照することとし、ここでは改めて説明はしない。

【0107】図からわかるように実施例 2 において、実施例 1 との違いは、第 3 の入力方向修正部 21 と、第 3 のビームフォーマ 22 と、第 2 の音声方向推定部 23、および有効雑音決定部 24 を追加した点である。

【0108】そして、第 2 及び第 3 のビームフォーマ 16、22 の出力、及び、雑音方向推定部 17 の出力、及び、第 1 及び第 2 の音声方向推定部 18、23 の出力を有効雑音決定部 24 に渡し、有効雑音決定部 24 の出力を第 1 の入力方向修正部 14 に渡す構成としてある。

【0109】このような構成の本システムの作用を説明する。まず、複数のマイクロホンを持つ音声入力部 11、この例では第 1 及び第 2 の計 2 本のマイクロホン m1、m2 を持つ音声入力部 11 で ch1、ch2 の音声を取り込む。そして、この音声入力部 11 から入力され

た 2 チャンネル分の音声の信号 ch1、ch2（すなわち、第 1 チャンネル ch1 は第 1 のマイクロホン m1 からの音声、第 2 チャンネル ch2 は第 2 のマイクロホン m2 からの音声に該当する）は、周波数分析部 12 に送られ、ここで例えば高速フーリエ変換（FFT）等の処理を行うことによって、それぞれのチャンネル別に周波数成分（周波数スペクトル）が求められる。

【0110】周波数分析部 12 でそれぞれ求められたチャンネル別の周波数成分は、それぞれ第 1、第 2 及び第 3 のビームフォーマ 13、16、22 に与えられる。

【0111】第 1 のビームフォーマ 13 では、2 チャンネル分の周波数成分入力について、目的音の方向対応に位相を合わせた上で、周波数領域の適応フィルタにより上述のようにして処理することで雑音を抑圧し、目的音の方向の周波数成分を出力する。ここで、具体的に説明すると第 1 の入力方向修正部 14 は第 1 のビームフォーマ 13 に対して次のような角度情報（ α ）を与える。つまり、第 1 の入力方向修正部 14 は、有効雑音決定部 24 を介して与えられる音声方向推定部 18 若しくは音声方向推定部 23 からの出力を用い、目的音の方向があたかもマイクロホンの正面方向となるよう、上記 2 チャンネルの周波数成分の入力位相を整えるに必要な角度情報（ α ）を入力方向修正量として第 1 のビームフォーマ 13 に対して与える。

【0112】この結果、第 1 のビームフォーマ 13 はこの修正量（ α ）対応に目的音方向を修正し、当該目的音方向以外の方向から到来する音声を抑圧させるようにすることで、雑音成分を抑圧し、目的音を抽出する。

【0113】つまり、第 2 および第 3 のビームフォーマ 16、22 の場合、雑音が目的音であるから、雑音に位相を合わせている。その結果、第 2、第 3 のビームフォーマ 16、22 では話者の音源は雑音源として扱われ、各ビームフォーマの内蔵する適応フィルタは話者音源からの音を抽出する処理をすることになるので、当該第 2、第 3 のビームフォーマ 16、22 の適応フィルタのパラメータからは話者音源の方向を反映した情報が得られることになる。

【0114】従って、第 1 または第 2 の音声方向推定部 18 または 23 により、第 2 または第 3 のビームフォーマ 16 または 22 における適応フィルタのパラメータを用いて雑音源方向を知れば、それは目的音である話者音源の方向を反映させたものである。従って、第 1 または第 2 の音声方向推定部 18 または 23 により、第 2 または第 3 のビームフォーマ 16 または 22 における適応フィルタのパラメータを反映させた出力を出し、第 1 の入力方向修正部 14 でこの出力対応に入力方向修正量（ α ）を発生し、この修正量対応に第 1 のビームフォーマ 13 における目的音方向を修正すれば、第 1 のビームフォーマ 13 は当該目的音方向以外の方向から到来する音声を抑圧するので、この場合、話者音源からの成分を

抽出できることになる。

【0115】一方、第1のビームフォーマ13の適応フィルタでは雑音成分が抽出されるようにパラメータが制御されているので、このパラメータから雑音方向推定部17では、雑音方向を推定し、その情報を第2及び第3の入力方向修正部15、21と有効雑音決定部24に与えることになる。

【0116】そして、当該雑音方向推定部17からの出力を受けた第2の入力方向修正部15では、当該雑音方向推定部17からの出力対応に入力方向修正量(α)を発生し、この修正量対応に第2のビームフォーマ16における目的音方向を修正すれば、第2のビームフォーマ16は当該目的音方向以外の方向から到来する音声を抑圧するので、この場合、話者音源以外の成分である雑音成分を抽出できることになる。

【0117】このとき、第2のビームフォーマ16の適応フィルタでは目的音である話者音声成分が抽出されるようにパラメータが制御されているので、このパラメータから第1の音声方向推定部18では、話者音声方向を推定することができる。そして、第1の音声方向推定部18はその推定した情報を有効雑音決定部24に与える。

【0118】また、雑音方向推定部17からの出力が第3の入力方向修正部21にも与えられているが、これを受けた第3の入力方向修正部21では、当該雑音方向推定部17からの出力対応に入力方向修正量(α)を発生し、第3のビームフォーマ22に与える。これにより、第3のビームフォーマ22はこの与えられた修正量対応に、自己における目的音方向を修正する。

【0119】これにより、第3のビームフォーマ22は当該目的音方向以外の方向から到来する音声を抑圧するので、この場合、話者音源以外の成分、つまり、雑音成分を抽出できることになる。このとき、第3のビームフォーマ22の適応フィルタでは目的音である話者音声成分が抽出されるようにパラメータが制御されているので、このパラメータから第2の音声方向推定部23では、話者音声方向を推定できる。そして、この推定した情報は有効雑音決定部24に与えることになる。

【0120】有効雑音決定部24では、第1および第2の音声方向推定部18、23から与えられた話者音声方向の推定情報と、雑音方向推定部17から与えられた雑音方向の推定情報とをもとに、第2のビームフォーマ16と第3のビームフォーマ22のいずれが雑音を有効に追尾しているかを判断する。そして、この判断結果に基づき、有効に追尾していると判断した方のビームフォーマにおける適応フィルタのパラメータを第1の入力方向修正部14に与える。

【0121】そのため、第1の入力方向修正部14では、当該パラメータを反映させた出力を出し、第1の入力方向修正部14でこの出力対応に入力方向修正量

(α)を発生し、この修正量対応に第1のビームフォーマ13における目的音方向を修正するので、第1のビームフォーマ13は当該目的音方向以外の方向から到来する音声を抑圧することになって、この場合、話者音源からの成分を抽出でき、しかも、広く移動する雑音源からの雑音を対象とする場合に、その移動する雑音源を見失うことなく、確実にとらえて雑音除去することが可能となる。

【0122】すなわち、この実施例においては、話者の音声周波数成分の抽出用として第1のビームフォーマ13が設けてあり、また、雑音周波数成分の抽出用として第2および第3のビームフォーマ16、22が設けてある。そして、観測点から見て図6に示すように、話者が 0° 方向に位置していて $0^\circ \pm \theta$ の角度範囲で監視すれば良いとすると、当該話者の音声周波数成分を抽出するために設けた第1のビームフォーマ13の変化範囲 ϕ 1、すなわち、適応フィルタにおける感度を高くする方向についての 1° 刻み変化範囲はせいぜい

$$-\theta < \phi 1 < \theta$$

に設定してこの範囲でフィルタリングに用いることになる。この場合、雑音周波数成分を抽出するために設けた第2および第3のビームフォーマ16、22のうち、第2のビームフォーマ16の変化範囲 ϕ 2は

$$-180^\circ + \theta < \phi 2 < -\theta$$

そして、第3のビームフォーマ22の変化範囲 ϕ 3は $\theta < \phi 3 < 180^\circ - \theta$

に設定することになる。但し、 180° は中心点を介して 0° の対向位置、 $-\theta$ は 0° 位置から見て図における反時計方向回り、 $+$ は時計方向回りを示す。

【0123】故に、このようにすると、第2のビームフォーマ16と第3のビームフォーマ22は、目的音到来範囲 ϕ 1を挟んで各々別々の範囲から到来する雑音を追尾することになる。そのため、 ϕ 2の範囲にあった雑音源が ϕ 1の範囲を横切って ϕ 3の範囲に急に移動した場合でも、 ϕ 3の領域を持ち場とする第3のビームフォーマ22が当該移動して来た雑音源を直ちに捕えることができるため、雑音方向を見失うことがなくなる。

【0124】この構成の場合、第2のビームフォーマ16の出力と、第3のビームフォーマ22の出力の計2つの出力が、雑音の出力として得られるが、雑音方向推定部17の結果に基づき、有効雑音決定部24において、第2のビームフォーマ16と第3のビームフォーマ22のいずれが雑音を有効に追尾しているかを判断し、この判断結果に基づき、有効に追尾して方の出力を雑音成分として用いることになる。

【0125】＜実施例2における全体の処理の流れ＞以上の処理の全体の流れを図7に示しておく。この処理はフレーム毎に行われる。各ビームフォーマの変化範囲および入力方向の初期値を設定した後に（ステップS31）、第1のビームフォーマ13の処理を行い（ステッ

ブ S 3 2)、雑音方向を推定した後に(ステップ S 3 3)、該雑音方向を入力として有効雑音決定部 2 4において、雑音方向が $\phi 2$ にあるか、 $\phi 3$ にあるかの判定を実施し、第 2のビームフォーマ 1 6と第 3のビームフォーマ 2 2のどちらを選択するかを決定する(ステップ S 3 4)。

【0 1 2 6】そして、推定された雑音方向が第 2の入力方向修正部 1 5あるいは第 3の入力方向修正部 2 1のどちらかに送られ、雑音方向が修正され、選択されたビームフォーマの処理が実行される。

【0 1 2 7】すなわち、推定された雑音方向が $\phi 2$ の領域であれば雑音方向が第 2の入力方向修正部 1 5に送られ、雑音方向が修正され、第 2のビームフォーマ 1 6の処理が実行され、目的音方向が推定される(ステップ S 3 4, S 3 5, S 3 6, S 3 7)。また、推定された雑音方向が $\phi 3$ の領域であれば雑音方向が第 3の入力方向修正部 2 1に送られ、雑音方向が修正され、第 3のビームフォーマ 2 2の処理が実行され、目的音方向が推定される(ステップ S 3 4, S 3 8, S 3 9, S 4 0, S 4 1)。

【0 1 2 8】次に、選択されたビームフォーマにより推定された音声方向(目的音方向)が $\phi 1$ の範囲内かどうか判断され、範囲内の場合は、推定された音声方向が第 1のビームフォーマ 1 3の第 1の入力方向修正部 1 4に送られ、入力方向の修正が実行される(ステップ S 4 2, S 4 3)。範囲外の場合は修正処理が実行されず、次のフレームに対する処理に進む(ステップ S 4 2, S 3 1)。

【0 1 2 9】この処理がフレーム毎に行われ、音声および雑音方向を追尾しながら、雑音抑圧が行われる。

【0 1 3 0】このように、実施例 2は、話者の発声した音声少なくとも異なる 2箇所以上の位置で受音する音声入力手段と、前記受音位置に対応する音声信号のチャンネル毎に周波数分析を行って複数チャンネルの周波数成分を出力する周波数分析手段と、この周波数分析手段にて得られる前記複数チャンネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声以外の音声を抑圧する到来雑音抑圧処理を行い、目的音声成分を得る第 1のビームフォーマ処理手段と、前記周波数分析手段にて得られる前記複数チャンネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声を抑圧し、第 1の雑音成分を得る第 2のビームフォーマ処理手段と、

前記第 1のビームフォーマ処理手段で計算されるフィルタ係数から雑音方向を推定する雑音方向推定手段と、前記第 2のビームフォーマ処理手段で計算されるフィルタ係数から第 1の目的音方向を推定する第 1の目的音方向推定手段と、前記第 3の適応ビームフォーマ処理手段で計算されるフィルタ係数から第 2の目的音方向を推定する第 2の目的音方向推定手段と、前記第 1のビームフォーマにおいて入力対象とする目的音の到来方向である第 1の入力方向を、前記第 1の目的音方向推定手段で推定された第 1の目的音方向と、第 2の目的音方向推定手段で推定された第 2の目的音方向のいずれか一方または両方に基づいて逐次修正する第 1の入力方向修正手段と、前記雑音方向修正手段で推定された雑音方向が所定の第 1の範囲にある場合に、前記第 2のビームフォーマにおいて入力対象とする雑音の到来方向である第 2の入力方向を該雑音方向に基づいて逐次修正する第 2の入力方向修正手段と、前記雑音方向修正手段で推定された雑音方向が所定の第 2の範囲にある場合に、前記第 3のビームフォーマにおいて入力対象とする雑音の到来方向である第 3の入力方向を該雑音方向に基づいて逐次修正する第 3の入力方向修正手段と、前記雑音方向推定手段で推定された雑音方向が所定の第 1の範囲から到来したか所定の第 2の範囲から到来したかに基づいて前記第 1の出力雑音と前記第 2の出力雑音のいずれか一方を真の雑音出力と決定していずれか一方の雑音を出力すると同時に、第 1の音声方向推定手段と第 2の音声方向推定手段のいずれの推定結果が有効であるかを決定していずれか一方の音声方向推定結果を第 1の入力方向修正手段へ出力する有効雑音決定手段とを具備して構成したものである。

【0 1 3 1】そして、このような構成の場合、話者の発声した音声を異なる 2箇所以上の位置で音声入力手段は受音し、周波数分析手段では、これを前記受音位置に対応する音声信号のチャンネル毎に周波数分析して複数チャンネルの周波数成分を出力する。そして、第 1のビームフォーマ処理手段はこの周波数分析手段にて得られる前記複数チャンネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声以外の音声を抑圧する到来雑音抑圧処理を行い、目的音声成分を得、また、第 2のビームフォーマ処理手段は、前記周波数分析手段にて得られる前記複数チャンネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声を抑圧し、雑音成分を得る。そして、雑音方向推定手段は、前記第 1のビームフォーマ処理手段で計算されるフィルタ係数から雑音方向を推定し、目的音方向推定手段は、前記第 2のビームフォーマ処理手段で計算されるフィルタ係数から目的音方向を推定する。また、第 1の目的音方向推定手段は前記第 2のビームフォーマ処理手段で計算されるフィ

ルタ係数から第 1 の目的音方向を推定し、第 2 の目的音方向推定手段は、前記第 3 の適応ビームフォーマ処理手段で計算されるフィルタ係数から第 2 の目的音方向を推定する。

【0132】また、第 1 の入力方向修正手段は、前記第 1 のビームフォーマにおいて入力対象とする目的音の到来方向である第 1 の入力方向を、前記第 1 の目的音方向推定手段で推定された第 1 の目的音方向と、第 2 の目的音方向推定手段で推定された第 2 の目的音方向のいずれか一方または両方に基づいて逐次修正する。そして、第 2 の入力方向修正手段は、前記雑音方向修正手段で推定された雑音方向が所定の第 1 の範囲にある場合に、前記第 2 のビームフォーマにおいて入力対象とする雑音の到来方向である第 2 の入力方向を該雑音方向に基づいて逐次修正し、第 3 の入力方向修正手段は、前記雑音方向修正手段で推定された雑音方向が所定の第 2 の範囲にある場合に、前記第 3 のビームフォーマにおいて入力対象とする雑音の到来方向である第 3 の入力方向を該雑音方向に基づいて逐次修正する。従って、第 2 の入力方向修正手段の出力により第 2 の入力方向を修正される第 2 のビームフォーマは第 2 の入力方向以外から到来する成分を抑圧して残りの雑音成分を抽出することになり、また、第 3 の入力方向修正手段の出力により第 3 の入力方向を修正される第 3 のビームフォーマは第 3 の入力方向以外から到来する成分を抑圧して残りの雑音成分を抽出することになる。

【0133】そして、有効雑音決定手段は、前記雑音方向推定手段で推定された雑音方向が所定の第 1 の範囲から到来したか所定の第 2 の範囲から到来したかに基づいて前記第 1 の出力雑音と前記第 2 の出力雑音のいずれか一方を真の雑音出力と決定していずれか一方の雑音を出力すると同時に、第 1 の音声方向推定手段と第 2 の音声方向推定手段のいずれの推定結果が有効であるかを決定して有効な方の音声方向推定結果を第 1 の入力方向修正手段へ出力する。この結果、目的音方向修正手段は、前記第 1 のビームフォーマにおいて入力対象となる目的音の到来方向である第 1 の入力方向を、前記決定した方の目的音方向推定手段で得た目的音方向に基づいて逐次修正するので、第 1 のビームフォーマは第 1 の入力方向以外から到来する雑音成分を抑圧して話者の音声成分を低雑音で抽出することになる。

【0134】このように本システムは雑音成分を抑圧した音声周波数成分と、音声成分を抑圧した雑音周波数成分とを別々に得ることができるが、この発明の最大の特徴は、第 1 乃至第 3 のビームフォーマとして、周波数領域で動作するビームフォーマを用いるようにした点にある。そして、このことによって、計算量を大幅に削減することができるようにしている。

【0135】そしてこの発明によると、適応フィルタの処理量が大幅に低減されるのに加え、入力音声に対する

周波数分析以外の周波数分析処理を省略することができ、かつ、フィルタ演算時に必要であった時間領域から周波数領域への変換処理も不要となり、全体の演算量を大幅に削減することができる。

【0136】また、本発明では、雑音追尾に監視領域を全く異ならせた雑音追尾用のビームフォーマを設けてあり、それぞれの出力からそれぞれ音声方向を推定させると共に、それぞれの推定結果からいずれが有効な雑音追尾をしているかを判断して、有効と判断された方のビームフォーマのフィルタ係数による音声方向の推定結果を第 1 の目的音方向修正手段に与えることで第 1 の目的音方向修正手段は、前記第 1 のビームフォーマにおいて入力対象となる目的音の到来方向である第 1 の入力方向を、前記目的音方向推定手段で推定された目的音方向に基づいて逐次修正するので、第 1 のビームフォーマは第 1 の入力方向以外から到来する雑音成分を抑圧して話者の音声成分を低雑音で抽出することができ、雑音源が移動してもこれを見失うことなく追尾して抑圧することができるようになるものである。

【0137】従来技術においては、2ch、すなわち、2本のマイクロホンだけでも目的音源の追尾を可能とすべく、雑音追尾用のビームフォーマを雑音抑圧のビームフォーマとは別に 1 個用いるが、例えば、雑音源が目的音の方向を横切って移動したような場合、雑音の追尾精度が低下することがあった。

【0138】しかし、本発明では、雑音を追尾するビームフォーマを複数用いて各々別個の追尾範囲を受け持つようにしたことにより、上記のような場合でも追尾精度の低下を抑止できるようになる。

【0139】以上の実施例 1 及び実施例 2 のシステムは、演算負荷の軽減を図りつつ、主として方向を持つ雑音について抑圧できるようにした例を示した。そして、この場合、テレビ会議システムなどのように、話者音源の配置がわかっていて、しかも、環境的に雑音が少ないような環境下での利用に適しているが、レベルも特性もまちまちで雑多な雑音の影響を受ける屋外や、大勢の人の集まる店舗や駅と云った所で使用するには不十分であると考えられる。

【0140】そこで、方向性の無い背景雑音も効果的に抑制できるようにした実施例を次に説明する。

【0141】(実施例 3) この実施例 3 は本発明の請求項 3 に対応する。ここでは、方向性のある雑音はビームフォーマにより抑圧し、方向性のない背景雑音はスペクトルサブトラクション (SS) 処理により、抑圧するようにした高精度の雑音抑圧が可能なシステムを説明する。

【0142】実施例 3 のシステムは、図 1 または図 5 の構成のシステムの後段に、更に図 8 の構成のスペクトルサブトラクション (SS) 処理部 30 を接続して構成する。スペクトルサブトラクション (SS) 処理部 30 は

図に示すように、音声帯域パワー計算部 31、雑音帯域パワー計算部 32、帯域重み計算部 33、スペクトル減算部 34 から構成されている。

【0143】これらのうち、音声帯域パワー計算部 31 は、前記ビームフォーマ 13 により得られた音声周波数を、周波数帯域毎に分割して帯域毎の音声パワーを計算するものであり、雑音帯域パワー計算部 32 は、前記ビームフォーマ 16 により得られた雑音周波数成分（またはビームフォーマ 16、22 によりそれぞれ得られ、有効雑音決定部 24 により選択されて出力された雑音周波数成分）を、周波数帯域毎に分割して帯域毎の雑音パワーを計算するものである。

【0144】帯域重み計算部 33 は、帯域 k 毎に、得られた音声の平均帯域パワー $P_v(k)$ と雑音の平均帯域パワー $P_n(k)$ を用い、帯域毎の帯域重み係数 $W(k)$ を計算するものであり、修正スペクトル減算部 34 は、前記入力帯域パワー計算部 31 にて計算された入力帯域パワーと、音声帯域パワー計算部 31 で計算された音声帯域パワーとに基き、音声信号の周波数帯域毎に重みをかけて背景雑音を抑圧するものである。

【0145】音声帯域パワー計算部 31 で用いる音声周波数成分と、雑音帯域パワー計算部 32 で用いる雑音周波数成分は、いずれも実施例 1 あるいは実施例 2 のビームフォーマの 2 つの出力である目的音声成分と雑音成分を利用する。そして、一般に、スペクトルサブトラクション (SS) として知られる雑音抑圧処理により、方向性のない背景雑音成分の抑圧を行う。

【0146】一般的に行われるスペクトルサブトラクション (SS) は、1 チャンネルのマイクロホン（つまり、1 本のマイクロホン）を用い、このマイクロホンの出力から音声のない区間において雑音のパワーを推定するため、非定常な雑音が音声に重畳している場合には対処できない。

【0147】また、2 チャンネルのマイクロホン（つまり、2 本のマイクロホン）を用いて、一方を雑音収集用、片方を雑音重畳音声収集用とする場合にも、両マイクロホンの設置場所を離す必要があり、その結果、音声に重畳する雑音と、雑音収集用マイクロホンで取り込む雑音との位相がずれ、スペクトルサブトラクションしても雑音抑圧の改善効果は大きく上がらなかった。

【0148】本実施例では、雑音成分を取り出すビームフォーマを用意して、このビームフォーマの出力を用いるようにしたため、実施例 1 および実施例 2 で述べたように、位相のずれが補正され、非定常雑音の場合でも高精度なスペクトルサブトラクション (SS) を実現できる。

【0149】さらに、周波数領域のビームフォーマの出力を利用しているため、周波数分析を省略してスペクトルサブトラクションが可能であり、従来より少ない演算量で非定常雑音を抑圧できる。

【0150】以下、具体的なスペクトルサブトラクション (SS) 方法について述べる。

【0151】＜スペクトルサブトラクション (SS) の原理＞まず、スペクトルサブトラクションの原理について説明する。目的音声用ビームフォーマ（第 1 のビームフォーマ 13）の出力を P_v 、雑音用ビームフォーマ（第 2 または第 3 のビームフォーマ 16 または 22）の出力を P_n とすると、

$$P_v = V + B'$$

$$10 \quad P_n = N + B''$$

と表すことができる。ここで、 V は音声成分のパワー、 B' は音声出力に含まれる背景雑音のパワーであり、 N は雑音源成分のパワー、 B'' は雑音出力に含まれる背景雑音のパワーである。これらのうち、音声出力成分に含まれる背景雑音成分を、スペクトルサブトラクション処理により抑圧する。

【0152】音声出力成分中の B' は、雑音出力成分中の B'' と同等であり、雑音源成分のパワー N も音声成分のパワー V に比べて小さいとすると、 $B' = P_n$ と考えることができ、スペクトルサブトラクション (SS) 処理用の重み係数 W は以下のように求めることができる。

すなわち、 W は

$$W = (P_v - P_n) / P_v \simeq V / (V + B')$$

となり、

$$V \simeq P_v * W$$

として音声成分を近似的に求めることができる。

【0153】図 8 にスペクトルサブトラクション (SS) 処理に必要な構成を、また、図 9 にスペクトルサブトラクション処理手順を示す。

30 【0154】2 つのビームフォーマ 13、15（または 22）からの出力として音声周波数成分と雑音周波数成分が得られる。ビームフォーマ 13 からの出力である音声周波数成分を用いて音声帯域パワー計算が実施され（ステップ S51）、ビームフォーマ 15（または 22）からの出力である雑音周波数成分を用いて雑音帯域パワー計算が実施される（ステップ S52）。ここでのパワー計算は、実施例 1 および実施例 2 で説明した本発明システムの音声周波数成分と雑音周波数成分を利用しており、これらはビームフォーマの処理を周波数領域で行っていることから、周波数分析なしに、そのまま音声および雑音の周波数成分の各帯域毎にパワーの計算を実行できる。

【0155】次に、計算されたパワー値を時間方向に平均化し、帯域毎に平均パワーを求める（ステップ S53）。帯域重み計算部 33 では、帯域 k 毎に、得られた音声の平均帯域パワー $P_v(k)$ と雑音の平均帯域パワー $P_n(k)$ を用い、次式により、帯域毎の帯域重み係数 $W(k)$ を計算する。

【0156】

$$W(k) = (P_v(k) - P_n(k)) / P_v(k) \quad (P_v(k) > P_n(k) \text{ の時})$$

$$W(k) = W_{min} \quad (P_v(k) \leq P_n(k) \text{ の時})$$

帯域重みは最大値 1.0 と最小値 W_{min} の間の値をとり、 W_{min} の値は例えば “0.01” 等とする。

【0157】次にスペクトル減算部 24 では、帯域重み計算部 23 で計算された帯域毎の重み係数 $W(k)$ を用い、入力音声周波数成分 $P_v(k)$ に重みをかけ、雑音成分を抑圧した音声周波数成分 $P_v(k)'$ を求める (ステップ S54)。

$$【0158】P_v(k)' = P_v(k) * W(k)$$

こうして、方向のない背景雑音はスペクトルサブトラクション (SS) 処理により、抑圧され、方向を持つ雑音は前述のビームフォーマにより抑圧されて、結果的に高精度の雑音抑圧が可能となる。

【0159】以上、この実施例 3 によれば、前記実施例 1 または実施例 2 の音抑圧装置において得られた音声周波数成分と雑音周波数成分を用いるようにしたものであり、前記周波数帯域毎に分割して帯域毎の音声パワーを計算する音声帯域パワー計算手段と、前記得られた雑音周波数成分を、周波数帯域毎に分割して帯域毎の雑音パワーを計算する雑音帯域パワー計算手段と、前記音声帯域パワー計算手段と雑音帯域パワー計算手段とから得られる音声と雑音の周波数帯域パワーに基き、音声信号の周波数帯域毎に重みをかけて背景雑音を抑圧するスペクトル減算手段とからなるスペクトル減算雑音抑圧手段を前記実施例 1 または実施例 2 の音抑圧装置にさらに具備して構成したものである。

【0160】この構成の場合、音声帯域パワー計算手段は、得られた音声周波数のスペクトル成分を、周波数帯域毎に分割して帯域毎の音声パワーを計算し、雑音帯域パワー計算手段は、前記得られた雑音周波数のスペクトル成分を、周波数帯域毎に分割して帯域毎の雑音パワーを計算する。そして、スペクトル減算手段は、前記音声帯域パワー計算手段と雑音帯域パワー計算手段とから得られる音声と雑音の周波数帯域パワーに基き、音声信号の周波数帯域毎に重みをかけて背景雑音を抑圧する。

【0161】この構成によれば、ビームフォーマでは抑圧できない方向性のない雑音 (背景雑音) は、本発明システムのビームフォーマで得ることのできる目的音声成分と雑音成分を利用し、これをスペクトルサブトラクション処理することで抑圧する。すなわち、本システムでは、ビームフォーマとして目的音声成分抽出用と雑音成分抽出用の 2 つのビームフォーマを備えているが、これらのビームフォーマの出力である目的音声成分と雑音成分を利用してスペクトルサブトラクション処理することにより、方向性のない背景雑音成分の抑圧を行う。スペクトルサブトラクション (SS) 処理は雑音抑圧処理として知られるが、一般的に行われるスペクトルサブトラ

36

クション (SS) 処理は、1 チャンネルのマイクロホン (つまり、1 本のマイクロホン) を用い、このマイクロホンの出力から音声のない区間において雑音のパワーを推定するため、非定常な雑音が音声に重畳している場合には対処できない。また、2 チャンネルのマイクロホン (つまり、2 本のマイクロホン) を用いて、一方を雑音収集用、片方を雑音重畳音声収集用とする場合にも、両マイクロホンの設置場所を離す必要があり、その結果、音声に重畳する雑音と、雑音収集用マイクロホンで取り込む雑音との位相がずれ、スペクトルサブトラクション処理しても雑音抑圧の改善効果は大きく上がらない。

【0162】しかし、本発明では、雑音成分を取り出すビームフォーマを用意して、このビームフォーマの出力を用いるようにしたため、位相のずれは補正されており、従って、非定常雑音の場合でも高精度なスペクトルサブトラクション処理を実現できる。さらに、周波数領域のビームフォーマの出力を利用しているため、周波数分析を省略してスペクトルサブトラクションが可能であり、従来より少ない演算量で非定常雑音を抑圧できる。

【0163】次に、実施例 3 を更に高精度化することができるようにした例を実施例 4 として次に説明する。

【0164】(実施例 4) 本実施例 4 は本発明の請求項 4 に対応する。本実施例は、実施例 3 のスペクトルサブトラクション (SS) において、雑音成分のパワーを修正することにより、さらに高精度に雑音抑圧を行うことを可能とするものである。すなわち、実施例 3 では雑音源のパワー N が小さいという仮定をおいたため、スペクトルサブトラクション (SS) 処理を行うと雑音源の成分が音声に重畳している部分では歪みが大きくなる懸念が拭えないという問題がある。

【0165】そこで、ここでは入力信号のパワーを用いて実施例 3 のスペクトルサブトラクションの帯域重みの計算を修正するようにする。

【0166】まず、音声出力パワーを P_v 、音声成分のパワーを V 、音声出力に含まれる背景雑音パワーを B' 、雑音出力パワーを P_n 、雑音源成分のパワーを N 、雑音出力に含まれる背景雑音成分を B'' 、どの信号も抑圧されていない入力信号のパワーを P_x とすると、

$$P_x = V + N + B$$

$$P_v = V + B'$$

$$P_n = N + B''$$

ここで、ここで、 $B \simeq B' \simeq B''$ と仮定すると、真の背景雑音成分のパワー P_b は、

$$\begin{aligned} P_b &= P_v + P_n - P_x \\ &= V + B' + N + B'' - (V + N + B) \\ &= B' + B'' - B \\ &= B \end{aligned}$$

となる。この雑音パワーを用いたスペクトルサブトラクション (SS) の重みは、

$$W = (P_v - P_b) / P_v$$

$$= (P_x - P_n) / P_v$$

と計算でき、背景雑音が非定常でかつ、Nが大きい場合でも歪みの少ないSS処理を行うことができる。

【0167】本実施例の構成を図10に示し、処理の流れを図11に示す。図10中、31は音声帯域パワー計算部、32は雑音帯域パワー計算部、34はスペクトル減算部、35は入力信号帯域パワー計算部である。

【0168】これらのうち、音声帯域パワー計算部31は、前記ビームフォーマ13により得られた音声周波数を、周波数帯域毎に分割して帯域毎の音声パワーを計算するものであり、雑音帯域パワー計算部32は、前記ビームフォーマ16または22により得られ、有効雑音決定部24により選択されて出力された雑音周波数成分を、周波数帯域毎に分割して帯域毎の雑音パワーを計算するものである。

【0169】入力帯域パワー計算部35は、前記周波数分析部12から得られた入力信号の周波数スペクトル成分を周波数帯域毎に分割し、帯域毎の入力パワーを計算するものであり、スペクトル減算部34は、前記入力帯域パワー計算部35にて計算された入力帯域パワーと、音声帯域パワー計算部31で計算された音声帯域パワーと、雑音帯域パワー計算部32で計算された雑音帯域パワーとに基づき、音声信号の周波数帯域ごとに重みをかけて背景雑音を抑圧するものである。

【0170】図10に示す実施例4でのスペクトルサブトラクション(SS)部30の構成と、実施例3でのスペクトルサブトラクション(SS)部30の構成との差は、実施例4においては何も抑圧されていない入力信号の周波数成分を更に用いる点である。

【0171】この入力信号周波数成分について、入力信号帯域パワー計算部35では、ビームフォーマからの音声周波数成分あるいは雑音周波数成分と同様に、帯域ごとにパワーを計算する(ステップS61)。

【0172】また、実施例3と同様に、2つのビームフォーマ13、15(または22)からの出力として音声周波数成分と雑音周波数成分が与えられるので、音声帯域パワー計算部31ではビームフォーマ13からの出力である音声周波数成分を用いて音声帯域パワー計算を実施し(ステップS62)、雑音帯域パワー計算部32ではビームフォーマ15(または22)からの出力である雑音周波数成分を用いて雑音帯域パワー計算を実施する(ステップS63)。

【0173】そして、スペクトル減算部34において、上述したように重み係数を求めた後に、重み付けを行う(ステップS64、S65)。これにより、方向を持つ雑音成分および方向のない雑音成分を抑圧した歪みの少ない音声成分のみの抽出ができるようになる。

【0174】このように、実施例4は、上記実施例3の雑音抑圧装置において、音声入力手段から得られた入力信号を周波数分析した入力信号の周波数成分を周波数帯

域毎に分割し、帯域毎の入力パワーを計算する入力帯域パワー計算手段を設けて、スペクトル減算手段には、入力帯域パワーと音声帯域パワーと雑音帯域パワーとに基づき、音声信号の周波数帯域毎に重みをかけて背景雑音を抑圧する処理を実施させるように構成したことを特徴とするものである。

【0175】この構成の場合、音声帯域パワー計算手段は、得られた音声周波数のスペクトル成分を、周波数帯域毎に分割して帯域毎の音声パワーを計算し、雑音帯域パワー計算手段は、前記得られた雑音周波数のスペクトル成分を、周波数帯域毎に分割して帯域毎の雑音パワーを計算する。また、入力帯域パワー計算手段があり、この入力帯域パワー計算手段は、音声入力手段から得られた入力信号を周波数分析して得た入力音声の周波数スペクトル成分を受けて、これを周波数帯域毎に分割し、帯域毎の入力パワーを計算する。そして、スペクトル減算手段は、前記音声帯域パワー計算手段と雑音帯域パワー計算手段とから得られる音声と雑音の周波数帯域パワーに基づき、音声信号の周波数帯域毎に重みをかけて背景雑音を抑圧する。

【0176】この実施例4においては、実施例3の構成におけるスペクトルサブトラクション処理において、更に雑音成分についてそのパワーを修正するようにしたことにより、一層高精度に雑音抑圧を行うことを可能とするものである。すなわち、第3の発明では雑音源のパワーNが小さいという仮定をおいたため、スペクトルサブトラクション処理を行うと雑音源の成分が音声に重畳している部分では歪みが大きくなることが避けられないが、ここでは入力信号のパワーを用いて第3の発明でのスペクトルサブトラクション処理における帯域重みの計算を修正するようにした。これにより、方向を持つ雑音成分および方向のない雑音成分を抑圧した歪みの少ない音声成分のみの抽出ができるようになるものである。

【0177】以上、種々の実施例を説明したが、本発明は第1には、話者の発声した音声を少なくとも異なる2箇所以上の位置で受音する音声入力手段と、前記受音位置に対応する音声信号のチャンネル毎に周波数分析を行って複数チャンネルの周波数成分を出力する周波数分析手段と、この周波数分析手段にて得られる前記複数チャンネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声以外の音声を抑圧する到来雑音抑圧処理を行い、目的音声成分を得る第1のビームフォーマ処理手段と、前記周波数分析手段にて得られる前記複数チャンネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声を抑圧し、雑音成分を得る第2のビームフォーマ処理手段と、前記第1のビームフォーマ処理手段で計算されるフィルタ係数から雑音方向を推定する

雑音方向推定手段と、前記第2のビームフォーマ処理手段で計算されるフィルタ係数から目的音方向を推定する目的音方向推定手段と、前記第1のビームフォーマにおいて入力対象となる目的音の到来方向である第1の入力方向を、前記目的音方向推定手段で推定された目的音方向に基づいて逐次修正する目的音方向修正手段と、前記第2のビームフォーマにおいて入力対象とする雑音の到来方向である第2の入力方向を、前記雑音方向推定手段で推定された雑音方向に基づいて逐次修正する雑音方向修正手段とを具備して構成したものである。

【0178】このような構成の場合、話者の発声した音声と異なる2箇所以上の位置で音声入力手段は受音し、周波数分析手段では、これを前記受音位置に対応する音声信号のチャンネル毎に周波数分析して複数チャンネルの周波数成分を出力する。そして、第1のビームフォーマ処理手段はこの周波数分析手段にて得られる前記複数チャンネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声以外の音声を抑圧する到来雑音抑圧処理を行い、目的音声成分を得、また、第2のビームフォーマ処理手段は、前記周波数分析手段にて得られる前記複数チャンネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声を抑圧し、雑音成分を得る。そして、雑音方向推定手段は、前記第1のビームフォーマ処理手段で計算されるフィルタ係数から雑音方向を推定し、目的音方向推定手段は、前記第2のビームフォーマ処理手段で計算されるフィルタ係数から目的音方向を推定する。目的音方向修正手段は、前記第1のビームフォーマにおいて入力対象となる目的音の到来方向である第1の入力方向を、前記目的音方向推定手段で推定された目的音方向に基づいて逐次修正するので、第1のビームフォーマは第1の入力方向以外から到来する雑音成分を抑圧して話者の音声成分を低雑音で抽出することになる。また、雑音方向修正手段は、前記第2のビームフォーマにおいて入力対象とする雑音の到来方向である第2の入力方向を、前記雑音方向推定手段で推定された雑音方向に基づいて逐次修正するので、第2のビームフォーマは第2の入力方向以外から到来する成分を抑圧して話者の音声成分を抑圧した残りの雑音成分を抽出することになる。

【0179】このように本システムは雑音成分を抑圧した音声周波数成分と、音声成分を抑圧した雑音周波数成分とを別々に得ることができるが、この発明の第1の特徴は、第1及び第2のビームフォーマとして、周波数領域で動作するビームフォーマを用いるようにした点にある。そして、このことによって、計算量を大幅に削減することができるようにしている。そしてこの発明によると、適応フィルタの処理量が大幅に低減されるのに加

え、入力音声に対する周波数分析以外の周波数分析処理を省略することができ、かつ、フィルタ演算時に必要であった時間領域から周波数領域への変換処理も不要となり、全体の演算量を大幅に削減することができる。

【0180】すなわち、従来技術では、ビームフォーマで抑圧できない拡散性雑音の抑圧処理のために、スペクトルサブトラクション処理を、ビームフォーマ処理の後に行うようにしており、このスペクトルサブトラクション処理は周波数スペクトルを入力とするため、FFT

(高速フーリエ変換)などの周波数分析が従来必要であったが、周波数領域で動作するビームフォーマを用いると当該ビームフォーマからは周波数スペクトルが出力されるため、これをスペクトルサブトラクション処理に流用できるので、特別にスペクトルサブトラクション処理のためのFFTを実施する従来のFFT処理工程は省略することができる。故に、全体の演算量を大幅に削減することができる。

【0181】また、ビームフォーマのフィルタを用いた方向推定の際に必要であった時間領域から周波数領域への変換処理も不要となり、全体の演算量を大幅に削減することができる。

【0182】また、第2には本発明は、話者の発声した音声と少なくとも異なる2箇所以上の位置で受音する音声入力手段と、前記受音位置に対応する音声信号のチャンネル毎に周波数分析を行って複数チャンネルの周波数成分を出力する周波数分析手段と、この周波数分析手段にて得られる前記複数チャンネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声以外の音声を抑圧する到来雑音抑圧処理を行い、目的音声成分を得る第1のビームフォーマ処理手段と、前記周波数分析手段にて得られる前記複数チャンネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声を抑圧し、第1の雑音成分を得る第2のビームフォーマ処理手段と、前記周波数分析手段にて得られる前記複数チャンネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声を抑圧し、第2の雑音成分を得る第2のビームフォーマ処理手段と、前記第1のビームフォーマ処理手段で計算されるフィルタ係数から雑音方向を推定する雑音方向推定手段と、前記第2のビームフォーマ処理手段で計算されるフィルタ係数から第1の目的音方向を推定する第1の目的音方向推定手段と、前記第3の適応ビームフォーマ処理手段で計算されるフィルタ係数から第2の目的音方向を推定する第2の目的音方向推定手段と、前記第1のビームフォーマにおいて入力対象とする目的音の到来方向である第1の入力方向を、前記第1の目的音方向推定手段で推定

された第 1 の目的音方向と、第 2 の目的音方向推定手段で推定された第 2 の目的音方向のいずれか一方または両方に基づいて逐次修正する第 1 の入力方向修正手段と、前記雑音方向修正手段で推定された雑音方向が所定の第 1 の範囲にある場合に、前記第 2 のビームフォーマにおいて入力対象とする雑音の到来方向である第 2 の入力方向を該雑音方向に基づいて逐次修正する第 2 の入力方向修正手段と、前記雑音方向修正手段で推定された雑音方向が所定の第 2 の範囲にある場合に、前記第 3 のビームフォーマにおいて入力対象とする雑音の到来方向である第 3 の入力方向を該雑音方向に基づいて逐次修正する第 3 の入力方向修正手段と、前記雑音方向推定手段で推定された雑音方向が所定の第 1 の範囲から到来したか所定の第 2 の範囲から到来したかに基づいて前記第 1 の出力雑音と前記第 2 の出力雑音のいずれか一方を真の雑音出力と決定していずれか一方の雑音を出力すると同時に、第 1 の音声方向推定手段と第 2 の音声方向推定手段のいずれの推定結果が有効であるかを決定していずれか一方の音声方向推定結果を第 1 の入力方向修正手段へ出力する有効雑音決定手段とを具備して構成する。

【0183】この第 2 の構成の場合、話者の発声した音声を変る 2 箇所以上の位置で音声入力手段は受音し、周波数分析手段では、これを前記受音位置に対応する音声信号のチャンネル毎に周波数分析して複数チャンネルの周波数成分を出力する。そして、第 1 のビームフォーマ処理手段はこの周波数分析手段にて得られる前記複数チャンネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声以外の音声を抑圧する到来雑音抑圧処理を行い、目的音声成分を得、また、第 2 のビームフォーマ処理手段は、前記周波数分析手段にて得られる前記複数チャンネルの周波数成分について、所望方向外の感度が低くなるように計算したフィルタ係数を用いての適応フィルタ処理を施すことにより前記話者方向からの音声を抑圧し、雑音成分を得る。そして、雑音方向推定手段は、前記第 1 のビームフォーマ処理手段で計算されるフィルタ係数から雑音方向を推定し、目的音方向推定手段は、前記第 2 のビームフォーマ処理手段で計算されるフィルタ係数から目的音方向を推定する。

【0184】また、第 1 の目的音方向推定手段は前記第 2 のビームフォーマ処理手段で計算されるフィルタ係数から第 1 の目的音方向を推定し、第 2 の目的音方向推定手段は、前記第 3 の適応ビームフォーマ処理手段で計算されるフィルタ係数から第 2 の目的音方向を推定する。

【0185】第 1 の入力方向修正手段は、前記第 1 のビームフォーマにおいて入力対象とする目的音の到来方向である第 1 の入力方向を、前記第 1 の目的音方向推定手段で推定された第 1 の目的音方向と、第 2 の目的音方向推定手段で推定された第 2 の目的音方向のいずれか一方

または両方に基づいて逐次修正する。そして、第 2 の入力方向修正手段は、前記雑音方向修正手段で推定された雑音方向が所定の第 1 の範囲にある場合に、前記第 2 のビームフォーマにおいて入力対象とする雑音の到来方向である第 2 の入力方向を該雑音方向に基づいて逐次修正し、第 3 の入力方向修正手段は、前記雑音方向修正手段で推定された雑音方向が所定の第 2 の範囲にある場合に、前記第 3 のビームフォーマにおいて入力対象とする雑音の到来方向である第 3 の入力方向を該雑音方向に基づいて逐次修正する。従って、第 2 の入力方向修正手段の出力により第 2 の入力方向を修正される第 2 のビームフォーマは第 2 の入力方向以外から到来する成分を抑圧して残りの雑音成分を抽出することになり、また、第 3 の入力方向修正手段の出力により第 3 の入力方向を修正される第 3 のビームフォーマは第 3 の入力方向以外から到来する成分を抑圧して残りの雑音成分を抽出することになる。

【0186】そして、有効雑音決定手段は、前記雑音方向推定手段で推定された雑音方向が所定の第 1 の範囲から到来したか所定の第 2 の範囲から到来したかに基づいて前記第 1 の出力雑音と前記第 2 の出力雑音のいずれか一方を真の雑音出力と決定していずれか一方の雑音を出力すると同時に、第 1 の音声方向推定手段と第 2 の音声方向推定手段のいずれの推定結果が有効であるかを決定して有効な方の音声方向推定結果を第 1 の入力方向修正手段へ出力する。この結果、目的音方向修正手段は、前記第 1 のビームフォーマにおいて入力対象となる目的音の到来方向である第 1 の入力方向を、前記決定した方の目的音方向推定手段で得た目的音方向に基づいて逐次修正するので、第 1 のビームフォーマは第 1 の入力方向以外から到来する雑音成分を抑圧して話者の音声成分を低雑音で抽出することになる。

【0187】このように本システムは雑音成分を抑圧した音声周波数成分と、音声成分を抑圧した雑音周波数成分とを別々に得ることができるが、この発明の最大の特徴は、第 1 及び第 2 のビームフォーマとして、周波数領域で動作するビームフォーマを用いるようにした点にある。そして、このことによって、計算量を大幅に削減することができるようにしている。

【0188】そしてこの発明によると、適応フィルタの処理量が大幅に低減されるのに加え、入力音声に対する周波数分析以外の周波数分析処理を省略することができ、かつ、フィルタ演算時に必要であった時間領域から周波数領域への変換処理も不要となり、全体の演算量を大幅に削減することができる。

【0189】また、本発明では、雑音追尾に監視領域を全く異ならせた雑音追尾用のビームフォーマを設けてあり、それぞれの出力からそれぞれ音声方向を推定させると共に、それぞれの推定結果からいずれが有効な雑音追尾をしているかを判断して、有効と判断された方のビー

ムフォーマのフィルタ係数による音声方向の推定結果を第1の目的音方向修正手段に与えることで第1の目的音方向修正手段は、前記第1のビームフォーマにおいて入力対象となる目的音の到来方向である第1の入力方向を、前記目的音方向推定手段で推定された目的音方向に基づいて逐次修正するので、第1のビームフォーマは第1の入力方向以外から到来する雑音成分を抑圧して話者の音声成分を低雑音で抽出することができ、雑音源が移動してもこれを見失うことなく追尾して抑圧することができるようになるものである。

【0190】従来技術においては、2ch、すなわち、2本のマイクロホンだけでも目的音源の追尾を可能とすべく、雑音追尾用のビームフォーマを雑音抑圧のビームフォーマとは別に1個用いるが、例えば、雑音源が目的音の方向を横切って移動したような場合、雑音の追尾精度が低下することがあった。

【0191】しかし、本発明では、雑音を追尾するビームフォーマを複数用いて各々別個の追尾範囲を受け持つようにしたことにより、上記のような場合でも追尾精度の低下を抑止できるようになる。

【0192】更に第3には、本発明は、上記第1または第2の音抑圧装置において、前記得られた音声周波数を、周波数帯域毎に分割して帯域毎の音声パワーを計算する音声帯域パワー計算手段と、前記得られた雑音周波数成分を、周波数帯域毎に分割して帯域毎の雑音パワーを計算する雑音帯域パワー計算手段と、前記音声帯域パワー計算手段と雑音帯域パワー計算手段とから得られる音声と雑音の周波数帯域パワーに基き、音声信号の周波数帯域毎に重みをかけて背景雑音を抑圧するスペクトル減算手段とからなるスペクトル減算雑音抑圧手段をさらに具備することを特徴とする。

【0193】この構成の場合、音声帯域パワー計算手段は、得られた音声周波数のスペクトル成分を、周波数帯域毎に分割して帯域毎の音声パワーを計算し、雑音帯域パワー計算手段は、前記得られた雑音周波数のスペクトル成分を、周波数帯域毎に分割して帯域毎の雑音パワーを計算する。そして、スペクトル減算手段は、前記音声帯域パワー計算手段と雑音帯域パワー計算手段とから得られる音声と雑音の周波数帯域パワーに基き、音声信号の周波数帯域毎に重みをかけて背景雑音を抑圧する。

【0194】この構成によれば、ビームフォーマでは抑圧できない方向性のない雑音（背景雑音）は、本発明システムのビームフォーマで得ることのできる目的音声成分と雑音成分を利用し、これをスペクトルサブトラクション処理することで抑圧する。すなわち、本システムでは、ビームフォーマとして目的音声成分抽出用と雑音成分抽出用の2つのビームフォーマを備えているが、これらのビームフォーマの出力である目的音声成分と雑音成分を利用してスペクトルサブトラクション処理することにより、方向性のない背景雑音成分の抑圧を行う。スペ

クトルサブトラクション（SS）処理は雑音抑圧処理として知られるが、一般的に行われるスペクトルサブトラクション（SS）処理は、1チャンネルのマイクロホン（つまり、1本のマイクロホン）を用い、このマイクロホンの出力から音声のない区間において雑音のパワーを推定するため、非定常な雑音が音声に重畳している場合には対処できない。また、2チャンネルのマイクロホン（つまり、2本のマイクロホン）を用いて、一方を雑音収集用、片方を雑音重畳音声収集用とする場合にも、両マイクロホンの設置場所を離す必要があり、その結果、音声に重畳する雑音と、雑音収集用マイクロホンで取り込む雑音との位相がずれ、スペクトルサブトラクション処理しても雑音抑圧の改善効果は大きく上がらない。

【0195】しかし、本発明では、雑音成分を取り出すビームフォーマを用意して、このビームフォーマの出力を用いるようにしたため、位相のずれは補正されており、従って、非定常雑音の場合でも高精度なスペクトルサブトラクション処理を実現できる。さらに、周波数領域のビームフォーマの出力を利用しているため、周波数分析を省略してスペクトルサブトラクションが可能であり、従来より少ない演算量で非定常雑音を抑圧できる。

【0196】更に第4には、本発明は、上記第3の発明の雑音抑圧装置において、音声入力手段から得られた入力信号を周波数分析した入力信号の周波数成分を周波数帯域毎に分割し、帯域毎の入力パワーを計算する入力帯域パワー計算手段を設けて、スペクトル減算手段には、入力帯域パワーと音声帯域パワーと雑音帯域パワーとに基き、音声信号の周波数帯域毎に重みをかけて背景雑音を抑圧する処理を実施させるようにすることを特徴とする。

【0197】この構成の場合、音声帯域パワー計算手段は、得られた音声周波数のスペクトル成分を、周波数帯域毎に分割して帯域毎の音声パワーを計算し、雑音帯域パワー計算手段は、前記得られた雑音周波数のスペクトル成分を、周波数帯域毎に分割して帯域毎の雑音パワーを計算する。また、入力帯域パワー計算手段があり、この入力帯域パワー計算手段は、音声入力手段から得られた入力信号を周波数分析して得た入力音声の周波数スペクトル成分を受けて、これを周波数帯域毎に分割し、帯域毎の入力パワーを計算する。そして、スペクトル減算手段は、前記音声帯域パワー計算手段と雑音帯域パワー計算手段とから得られる音声と雑音の周波数帯域パワーに基き、音声信号の周波数帯域毎に重みをかけて背景雑音を抑圧する。

【0198】この第4の発明においては、第3の発明のスペクトルサブトラクション（SS）処理において、更に雑音成分についてそのパワーを修正するようにしたことにより、一層高精度に雑音抑圧を行うことを可能とするものである。すなわち、第3の発明では雑音源のパワーNが小さいという仮定をおいたため、スペクトルサブ

トラクション (SS) 処理を行うと雑音源の成分が音声に重畳している部分では歪みが大きくなることが避けられないが、ここでは入力信号のパワーを用いて第3の発明でのスペクトルサブトラクション処理における帯域重みの計算を修正するようにした。これにより、方向を持つ雑音成分および方向のない雑音成分を抑圧した歪みの少ない音声成分のみの抽出ができるようになるものである。

【0199】尚、本発明は上述した実施例に限定されるものではなく、種々変形して実施可能である。

【0200】

【発明の効果】以上、詳述したように、本発明によれば、全体の演算量を大幅に削減することができ、また、ビームフォーマのフィルタを用いた方向推定の際に必要なであった時間領域から周波数領域への変換処理も不要となり、全体の演算量を大幅に削減することができると云う効果が得られる。

【0201】また、本発明では、雑音成分を取り出すビームフォーマを用意して、このビームフォーマの出力を用いるようにしたため、位相のずれは補正されており、従って、非定常雑音の場合でも高精度なスペクトルサブトラクション処理を実現できる。さらに、周波数領域のビームフォーマの出力を利用しているため、周波数分析を省略してスペクトルサブトラクションが可能であり、従来より少ない演算量で非定常雑音を抑圧できて、方向性のある雑音成分ばかりか、方向性のない雑音成分（背景雑音）も抑圧できて歪みの少ない音声成分の抽出ができるようになる云う効果が得られる。

【図面の簡単な説明】

【図1】本発明の実施例1の全体構成を示すブロック図である。

【図2】本発明で使用するビームフォーマの構成例と動作例を説明する図である。

【図3】本発明の実施例1における方向推定部の作用を説明するためのフローチャートである。

【図4】本発明の実施例1におけるシステムの作用を説明するためのフローチャートである。

【図5】本発明の実施例2の全体構成を示すブロック図である。

【図6】本発明の実施例2におけるビームフォーマの追尾範囲を説明するための図である。

【図7】本発明の実施例2におけるシステムの作用を説明するためのフローチャートである。

【図8】本発明の実施例3の要部構成を示すブロック図である。

【図9】本発明の実施例2におけるシステムの作用を説明するためのフローチャートである。

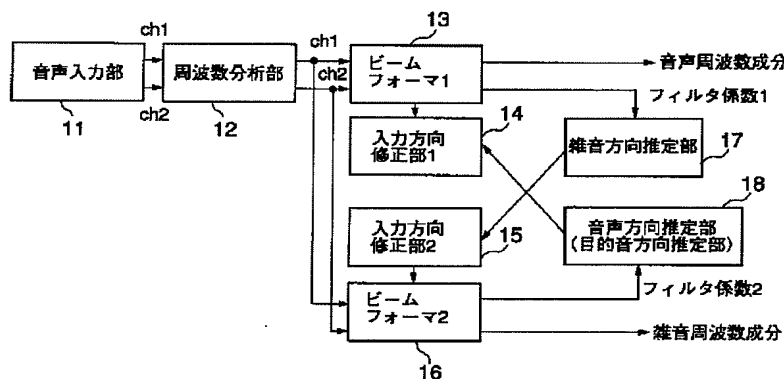
【図10】本発明の実施例4の要部構成を示すブロック図である。

【図11】本発明の実施例2におけるシステムの作用を説明するためのフローチャートである。

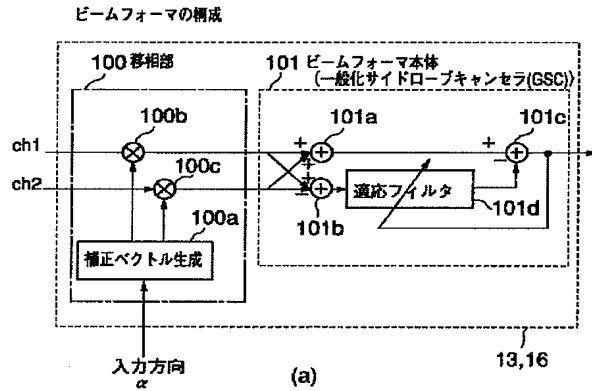
【符号の説明】

- 11…音声入力部
- 12…周波数解析部
- 13…第1のビームフォーマ
- 14…第1の入力方向修正部
- 15…第2の入力方向修正部
- 16…第2のビームフォーマ
- 17…雑音方向推定部
- 18…第1の音声方向推定部（目的音方向推定部）
- 21…第3の入力方向修正部
- 22…第3のビームフォーマ
- 23…第2の音声方向推定部
- 24…有効雑音決定部
- 30…スペクトルサブトラクション (SS) 処理部
- 31…音声帯域パワー計算部
- 32…雑音帯域パワー計算部
- 33…帯域重み計算部
- 34…スペクトル減算部
- 35…入力信号帯域パワー計算部

【図1】

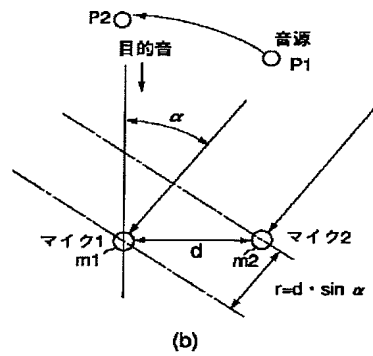
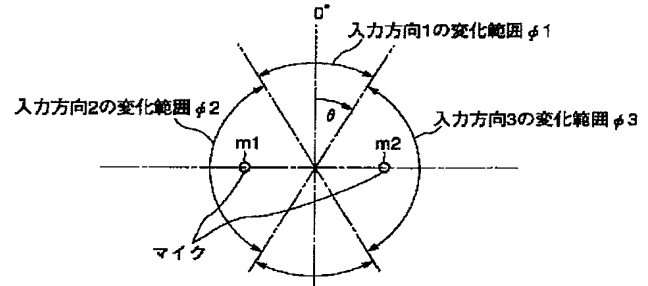


【図2】



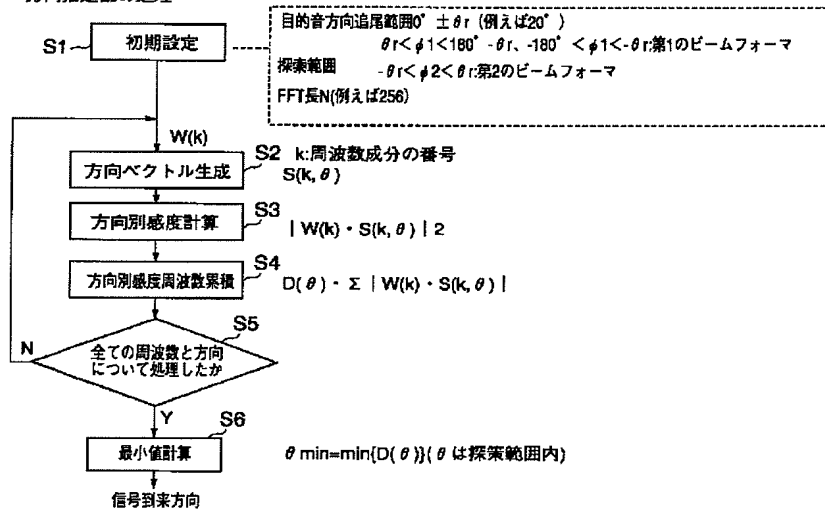
【図6】

〔ビームフォーマの追尾範囲〕

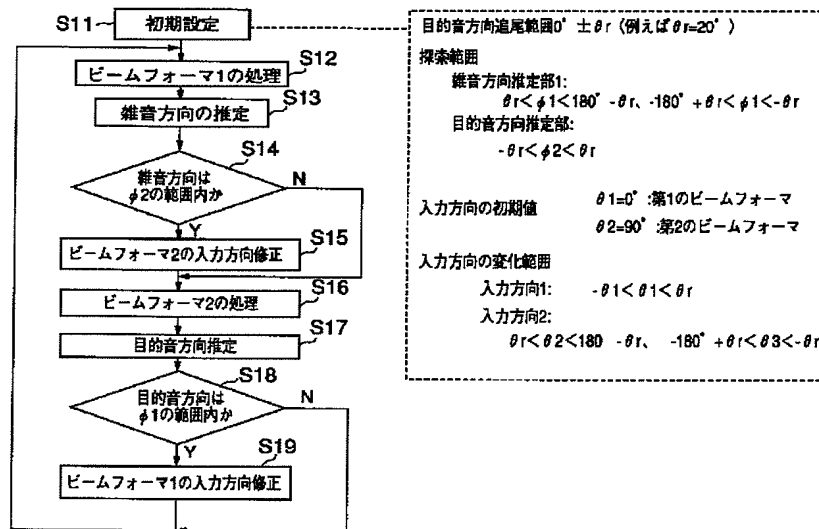


【図3】

方向推定部の処理

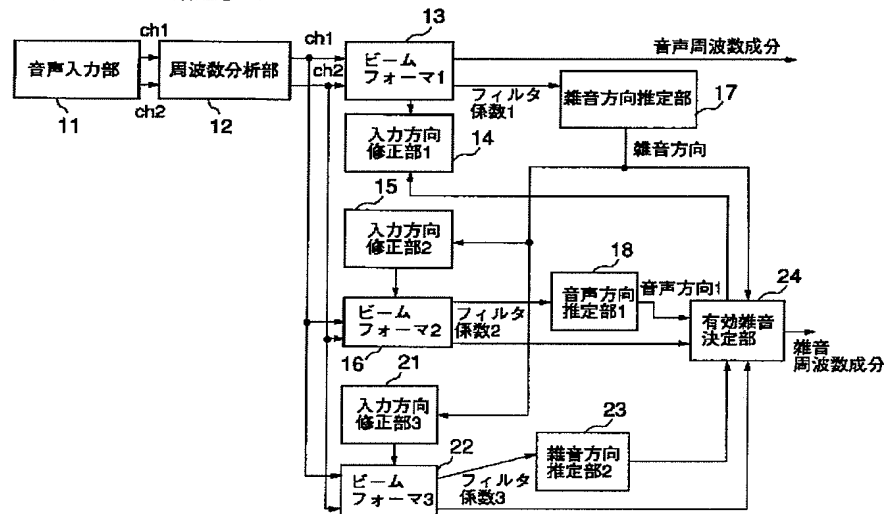


【図4】

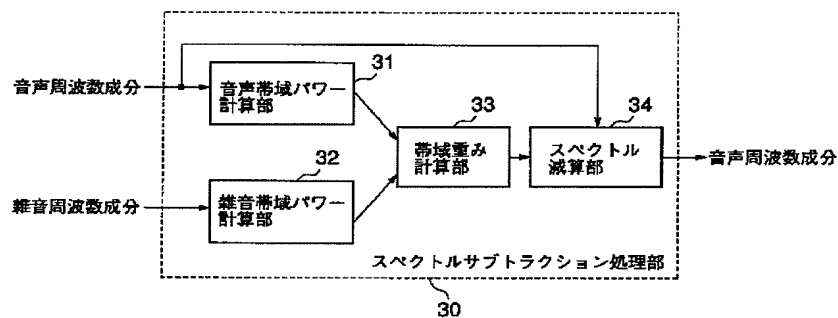


【図5】

2ch周波数領域BF+3BF話者追尾

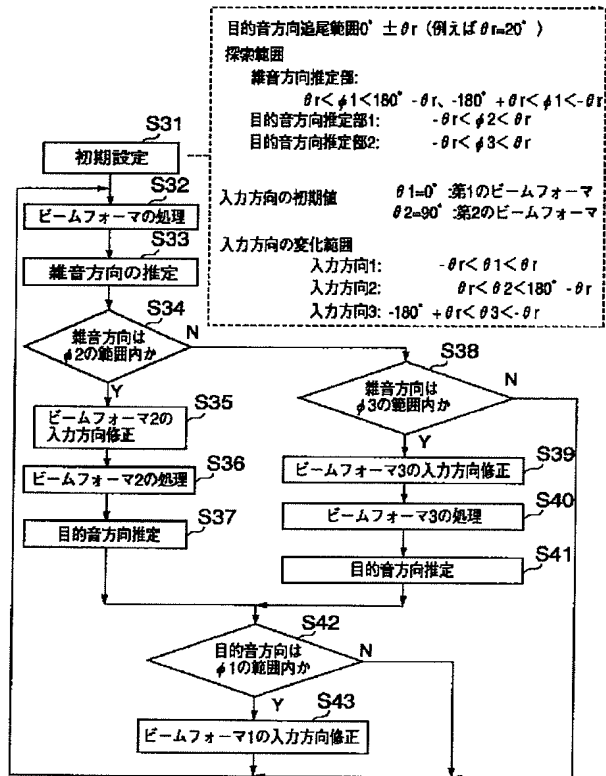


【図8】



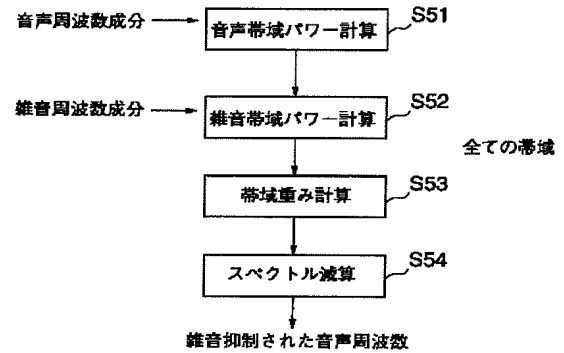
【図7】

[実施例2の処理の流れ]



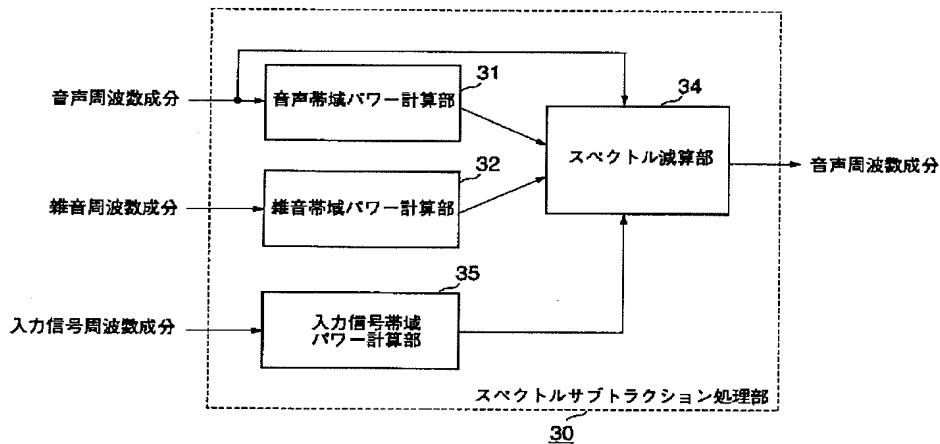
【図9】

2chSSの全体処理の流れ



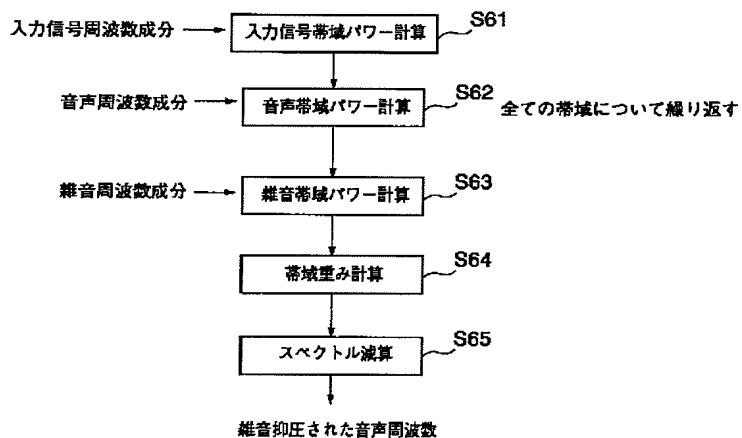
【図10】

修正2chSSの構成



【図11】

修正版2chSSの全体処理の流れ



フロントページの続き

(51) Int. Cl.⁷

識別記号

F I

ターマード (参考)

H 0 3 H 21/00

F ターム (参考) 5D015 CC02 CC14 DD02 EE05
 5J023 DA05 DB02 DC06 DC08 DD03
 5J083 AA05 AB10 AC07 AC15 AC18
 AC30 AD15 BC01 BE12 BE14
 BE18 BE43 BE53 BE58 CA10
 CA12